

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2005, 04:00:43 ; Search time 5145 Seconds
(without alignments)
10397.387 Million cell updates/sec

Title: US-10-642-531-1

Perfect score: 1104

Sequence: 1 atgcgcgtcgaagaagcctag.....cgcagagccttcgcatga 1104

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank:*

1: gb_ba:*

2: gb_hcg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pac:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_scs:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the local score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	878.4	79.6	1132	8	AY059083 Arabidops
2	878.4	79.6	1441	8	AF039406 Arabidops
3	878.4	79.6	1483	8	AY035017 Arabidops
4	873.6	79.1	1457	6	AR269261 Arabidops
5	873.6	79.1	1457	8	ATW3312 Arabidops
6	639.2	57.9	1537	8	AJ630654 Oryza sat
7	545.8	49.4	1500	8	AY026039 Oryza sat
8	542.2	49.1	1422	8	AK069157 Oryza sat
9	524.8	47.5	1424	8	BT009355 Triticum
10	524.6	47.5	1602	8	AF038586 Zea mays
11	517.4	46.3	1380	8	AF038585 Zea mays
12	511.4	46.3	1535	8	AK100033 Oryza sat
13	442	40.0	1405	8	AF33611 Oryza sat
14	324.2	29.4	899	6	AR566929 Sequence
15	324	29.3	86022	8	ATAC011623 Sequence
16	208.2	18.9	371	6	BD267332 Composite
17	208.2	18.9	371	6	AR566843 Sequence
18	191.8	17.4	110000	8	CR382129 Continuation (29 o
19	191.8	17.3	94704	8	AC020580 Arabidops

C 20	177.8	16.1	129306	8	NCB24N11	BX284751 Neurospor
C 21	177.8	16.1	199386	8	NCB14D6	AL561173 Neurospor
C 22	171.2	15.5	3326	8	AY589088	AY589088 Glycine m
C 23	163.8	14.8	110000	2	AP006500_05	Continuation (6 of
C 24	154.4	14.0	33415	8	AY589089	AY589089 Glycine m
C 25	137	12.4	109399	8	AP003749	AP003749 Oryza sat
C 26	137	12.4	139039	2	AP004383	AP004383 Oryza sat
C 27	135	12.2	111484	8	AC144475	AC144475 Medicago
C 28	135	12.2	111486	8	AC136140	AC136140 Medicago
C 29	135	12.2	111486	8	AC147536	AC147536 Medicago
C 30	135	12.2	135914	8	AC146757	AC146757 Medicago
C 31	130.8	11.8	94647	8	AC145372	AC145372 Medicago
C 32	130.8	11.8	94656	8	AC138579	AC138579 Medicago
C 33	121.4	11.0	236	6	AX645560	AX645560 Sequence
C 34	121.4	11.0	236	6	AX676721	AX676721 Sequence
C 35	120.2	10.9	624	6	AR549025	AR549025 Sequence
C 36	120.2	10.9	1536	6	AX489451	AX489451 Sequence
C 37	118.2	10.7	158714	8	AC082644	AC082644 Oryza sat
C 38	104.2	9.4	33733	8	SPAC644	AL35012 S.pombe c
C 39	104.2	9.4	110000	8	CR382135_11	Continuation (12 o
C 40	99.4	9.0	1149	10	RNU27456	U27456 Rattus norv
C 41	99.4	9.0	1654	10	AP043070	AP043070 Mus muscu
C 42	99.4	9.0	1822	10	BC046595	BC046595 Mus muscu
C 43	90.2	8.2	3212	5	BC070978	BC070978 Xenopus l
C 44	89.2	8.1	102197	8	CR382126_25	Continuation (26 o
C 45	86.8	7.9	110000	8	AE016818_11	Continuation (12 o

ALIGNMENTS

RESULT 1	AY059083	1132 bp	mRNA	linear	PLN 18-SEP-2002
LOCUS	AY059083				
DEFINITION	Arabidopsis thaliana putative pyruvate dehydrogenase kinase				
ACCESSION	(At3g06483) mRNA, complete cds.				
VERSION	AY059083.1	GI:16323389			
KEYWORDS	FLI CDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	1 (bases 1 to 1132)				
AUTHORS	Yamada, K., Banh, J., Sakano, H., Pham, P.K., Banh, J., Egu, P., Lee, J.M., Toriumi, M., Yu, G., Brooks, S., Chao, Q., Chen, H., Karlin-Neumann, G., Kim, C., Lam, B., Miranda, M., Nguyen, M., Palm, C.J., Shin, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.				
TITLE	Arabidopsis Open Reading Frame (ORF) Clones				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1132)				
AUTHORS	Yamada, K., Banh, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bower, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shin, P., Southwick, A., Tracy, S.E., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-OCT-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA				
COMMENT	The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFLV cDNAs (RFLV cDNA : RIKEN Arabidopsis Full-length cDNA). Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.				
	The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pLNT (ORF) clones using the RFLV cDNAs: Yamada, K., Banh, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M.,				

Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Katlin-Neumann, G., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Tacey, S.B., Davis, R.W., Becker, J.R. and Theologis, A.

Yamada, K. (SSP/PEEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP/PEEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES

source

Location/Qualifiers

1..1132

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="3"

/clone="U10102"

/ecotype="Columbia"

/note="This clone is in pUNI 51."

1..1132

/gene="At3g06483"

1..1101

/gene="At3g06483"

/codon_start=1

/evidence=experimental

/product="putative pyruvate dehydrogenase kinase"

/protein_id="AA15189.1"

/db_xref="GI:16323390"

/translation="MAVKKACEMPEKSLIEDVHKCMKQTVSLVYMEFSGKPTER
NLIISQFLHKELEIRVARRAIELQTLPGYLSIDKPAVLKRDVLYLSPDMRFPK
DSGDEDTOMIKAVKRVNNVPMALGVNOLKGNNSGNDLHIOFDRFLSISG
IRMLIOGHVELHNPNPLHTVGYIHTKMSMEVARNASBARSICREXGSAENIY
GDSFPEPVYPTLHLMYELVNKSLRANRERPVDSRVAAPRIITVAGIEPVTVK
SDSGGIARSGLPRIETVLYASTARNPLEEVDGIDVPVTAAGVIGLPISLVARY
FGSDLIISMEGIGIDVYHLSRLSDSPELP"

1102..1132

/gene="At3g06483"

ORIGIN

3'UTR

Query Match 79.6%; Score 878.4; DB 8; Length 1132;
Best Local Similarity 87.9%; Pred. No. 1.3e-248;
Matches 970; Conservative 0; Mismatches 131; Indels 3; Gaps 1;

1 ATGGCGGTGAAGAAGCTACGAGATGTTTTCGAGAAGCTTGATCGAGACGTTTCACAGA 60
1 ATGGCASTGAAGAAGCGCTGCGAAATGTTCCGAGAAGTTGATTCGAAGATGTTCCAAA 60
61 TGGGAGTGCATGAAGCAGACGGCGCTGAGCTTCAAGTATCATGATGAGTTGGTTCCACT 120
61 TGGGAGTGCATGAAGCAGACGGCGCTGAGCTTCAAGTATCATGATGAGTTGGTTCCACT 120
121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAAGGTTCCGATTCCG 180
121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAAGGTTCCGATTCCG 180
121 CCGACTGAGAGAAATCTTTTGTATTTCTGCTCACTTTTTCATGAGAGGCTTCGATTCCG 180
181 ATGCGAGCGCTGCGATCGAAGCTCGAGACGCTGCTTATGAGCTTCTGAGAAACCTGCG 240
181 ATGCGAGCGCTGCGATCGAAGCTCGAGACGCTGCTTATGAGCTTCTGAGAAACCTGCG 240
181 GTCGCGAGAGAGCGATCGAAGCTCGAGACGCTTCTTATGAGCTTCTGATTAACCTGCG 240
241 GTCCTGAAGGTAAAGATGATGATGAGTCACTTCAAGGAGCATGAGAGCGTTTCTCTGAG 300
241 GTCCTGAAGGTAAAGATGATGATGAGTCACTTCAAGGAGCATGAGAGCGTTTCTCTGAG 300
301 ATCAAGATATCTGCTGATGAGAAAGAGTTCACACAGATGATCAAGGCTTTAAAGTAAGG 360
301 ATCAAGATATCTGCTGATGAGAAAGAGTTCACACAGATGATCAAGGCTTTAAAGTAAGG 360
301 ATTAAGATATCGGCTGAGAGAAAGATTTCACTCAGATATTAAGGCTTCAAAGTAAGG 360
361 CACAAACAGCTGTTCCCTCATGATGCTGCTGCTGCTGAGAACCACTGAGAAAGAAATGAA 420
361 CATAACATGTGTTCCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419

421 CTCTACGAAAAGCTTGATGATGATTCATCAGTTTCTTGATCGCTTCTACTGTCTCTATA 480
420 --TTCTGGAATCTTGATGATGATTCATCAGTTTCTTGATCGCTTCTACTGTCTCTATA 477
481 GGGATCCGATGCTTATCGGGGAGCATGTTGAGTTGATTCATTAACCAACCCACACTTGC 540
478 GGGATCCGATGCTTATCGGGGAGCATGTTGAGTTGATTCATTAACCAACCCACACTTGC 537
541 ACAGTGGTTACATACACACCAAGATGCTCTTATGAGAGTGGCAAGAACTGATGTA 600
538 ACAGTGGTTATATACACACCAAGATGCTCTTATGAGAGTGGCAAGAACTGATGTA 597
601 GATGCAAGTCCATTTGTTTTCAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
598 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
661 GATGCAAGTCCATTTGTTTTCAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
658 GATGCAAGTCCATTTGTTTTCAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
721 GTCAGAACTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
718 GTCAGAACTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777
781 CCAATCCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
778 CCAATCCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
841 GGTGAGGTATACCGAGAAAGCGCTCTCCCTAAATATTCATCTTACCTGATGAGCACTGCA 900
838 GGTGAGGTATACCGAGAAAGCGCTCTCCCTAAATATTCATCTTACCTGATGAGCACTGCA 897
901 AAAAACCACCTGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
898 AAAAACCACCTGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 957
961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
958 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017
1021 CAGATCATATTCATGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1018 CAGATCATATTCATGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
1081 GACTGCGAGAGCGCTTGGCCATGA 1104
1078 GATTGCGAGAGCGCTTGGCCATGA 1101

RESULT 2
AF039406 1441 bp mRNA linear PLN 22-DEC-1998
LOCUS
DEFINITION Arabidopsis thaliana pyruvate dehydrogenase kinase (PDK) mRNA,
nuclear gene encoding mitochondrial protein, complete cds.
ACCESSION AF039406
VERSION AF039406.1 GI:4049631
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi-
s
1 (bases 1 to 1441)
Thelen, J.J., Miernyk, J.A. and Randall, D.D.
Nucleotide and deduced amino acid sequences of the pyruvate
dehydrogenase kinase from Arabidopsis thaliana (Accession No.
AF039406) (F898-192)
Plant Physiol. 118 (4), 1533 (1998)
JOURNAL 2 (bases 1 to 1441)
REFERENCE Thelen, J.J., Miernyk, J.A. and Randall, D.D.
AUTHORS Direct Submission
TITLE Submitted (20-DEC-1997) Biochemistry, University of Missouri, 117
JOURNML Schweitzer Hall, Columbia, MO 65211, USA

The Saik, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Liu, S. X., Sakano, H., Pham, P. K., Bann, J., Chung, M. K., Goldsmith, A. D., Lee, J. M., Quach, H. L., Tang, C. C., Toriumi, M., Yu, G., Bower, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Kosem, E., Lam, B., Lin, J., Meyers, M. C., Miranda, M., Nguyen, M., Palm, C. J., Shinn, P., Southwick, A., Davis, R. W., Ecker, J. R. and Theologis, A.

Yamada, K. (SSP/PGSC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGSC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES

source
1. 1483
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAFL05-09-E15 (R10102)"
/ecotype="Columbia"
/note="This clone is in a modified pBluescript vector (lambda Zap) as a XhoI/SacI insert."
gene
1. 1483
/gene="At3g06483"
5' UTR
1. 143
/gene="At3g06483"
144. 1244
/gene="At3g06483"
/codon_start=1
/evidence="experimental"
/product="putative pyruvate dehydrogenase kinase"
/protein_id="AAK59522.1"
/db_xref="GI:14334688"
/translation="MAVKKACEMFPKSLIEDVHKCMKQTVSLRYMMEFSGKPTER
NLISAQFLHEKLEIRVARVAIEQLPYGLSDPAVLKVDVLYLESFDDMAFPIK
DSGDEKPTOMIKAVYKRNHNVPMALGVNQLKKNNGNLDIEHIFDRLFLSIG
IRMLIQHVELHNPNPLTVGYIHTKMSMEVARSASEPARICFREYSGAPENIY
GDSPTFPEVPTLHLMYELVKNLSLRVQDERVDSRVAIPRIIVAGIDIVITKV
SDGGGIGARSGELRIFTYVLTARNLEEDVDIGIADVVTNAGYRGIPISRLYARY
FGGDLIIISMEGISTDAYLHLSRLDSQEPPLP"
1245. 1483
/gene="At3g06483"
3' UTR
1467
/gene="At3g06483"
misc_feature
1467
/gene="At3g06483"
/note="artifact within poly A tail"

ORIGIN

Query Match 79.6%; Score 878.4; DB 8; Length 1483;
Best Local Similarity 87.9%; Pred. No. 1.4e-248;
Matches 970; Conservative 0; Mismatches 131; Indels 3; Gaps 1;
Db 1 ATGGCGGTGAAGAAGCTAGACGATGTTTTCGAGAAGCTGATCGAGACGTTACACA 60
144 ATGGCAGTGAAGAAGCTAGACGATGTTTTCGAGAAGCTGATCGAGAAGTTTACACA 203
QY 61 TGGGATGATGAAGAGAGACGCGCTGAGCTCAGGTATGATGATGAGTTCGGTTCAC 120
Db 204 TGGGGTTGATGAAGAGAGACGCGCTGAGCTCAGGTATGATGATGAGTTCGGTTCAC 263
QY 121 CCCACTGAGAGAACTTCTGATCTCGGCGCAGTTTCTTCAACAAGAGCTTCCGATTCG 180
Db 264 CTTACTGAGAGAACTTCTGATCTCGGCGCAGTTTCTTCAACAAGAGCTTCCGATTCG 323
QY 181 ATCGGAGCGCTGCGATCGAATCGAGAGCTGCTTATGCGCTCTGAGAGAACTCGCC 240
Db 324 GTGCGCAGAGAGCGATCGAATCGAGAGCTTCTTATGCTCTCTGATTAACCTCGC 383
QY 241 GTCTTGAAGTGAAGATTTGATGTGAGTCAATTCAGGACATGAGACGTTTCTTGAG 300
Db 384 GTTTGAAGGTGCGGATTTGATTTGATCTTCAAGGACATGAGACATTTCTCTGAG 443

QY 301 ATCAAGATAGTCTGATGAGAGAGAGTTCACACAGATGATCAAGGCTTTAAAGTAAG 360
Db 444 ATTAAGATTCGGGTACAGAGAGAGAGTTCACACAGATGATTAAGGCTTCAAGTAAG 503
QY 361 CACACAGAGTGTTCCTCATGATGCTGCGGTGAGAACCACTGAGAGAGAGTAAGAA 420
Db 504 CATACAGATGTGTTCCTCATGATGCTGCGGTGAGAACCACTGAGAGAGTAAGAA 562
QY 421 CTCTAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
Db 563 --TTCTGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 620
QY 481 GGGATCCGATGCTTATCGGAGAGAGTGTGATGATGATGATGATGATGATGATGATG 540
Db 621 GGGATCCGATGCTTATCGGAGAGAGTGTGATGATGATGATGATGATGATGATGATG 680
QY 541 ACAGTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 681 ACAGTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 740
QY 601 GATGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db 741 GATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 800
QY 661 GATCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 801 GATCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 860
QY 721 GTCAAGAACTCTCTCGTGTGATGATGATGATGATGATGATGATGATGATGATG 780
Db 861 GTCAAGAACTCTCTCGTGTGATGATGATGATGATGATGATGATGATGATGATG 920
QY 781 CCAATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db 921 CCAATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 980
QY 841 GGTGAGGTATGAG 900
Db 981 GGTGAGGTATGAG 1040
QY 901 AGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 960
Db 1041 AGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1100
QY 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db 1101 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1160
QY 1021 CAGATCATATCATGAG 1080
Db 1161 CAGATCATATCATGAG 1220
QY 1081 GACTCGAG 1104
Db 1221 GACTCGAG 1244

RESULT 4
AR269261
LOCUS AR269261
DEFINITION Sequence 1 from patent US 650670.
ACCESSION AR269261
VERSION AR269261.1
KEYWORDS GI:29700268
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1457)
AUTHORS Zou, J. and Taylor, D.C.
TITLE Plant pyruvate dehydrogenase kinase gene
JOURNAL Patent: US 650670-A 1-31-DEC-2002;
FEATURES Location/Qualifiers

source 1. .1457
/organism="unknown"
/mol_type="genomic DNA"

Query Match 79.1%; Score 873.6; DB 6; Length 1457;
Best Local Similarity 87.6%; Pred. No. 3.7e-247;
Matches 967; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

1 ATGGCGGTGAAGAGGCTAGCCGAGATGTTTCCGAAGCTTGAATCCAGAGCTTCACAGA 60
104 ATGGCAGTAAAGAAACCTCGGAAATGTTCCGAAGGTTGATCGAAGATGTTCAAAA 163
61 TGGGAGTACGTAAGAGAGACGGCGTGAAGCTCAGGATCAATGATGAGATTCGATTCAC 120
164 TGGGGTTCATGAGCAAAACCGGTGTAAGCTTATGATATGATGATGATGATGATGATG 223
121 CCCACTGAGAGAAACCTTCTGATCTGGCGCAGTTTCTTCAAGAGCTTCGATTCGG 180
224 CCGTACTGAGAGAAATCTTTGATTTTCTGCTCAGTTTTCATAGAGAGCTTCGATTCGG 283
181 ATGCGAGGCGTCCGATTCGAACTCGAGAGCGCTGCTTATGAGCTCTCTGAGAAACCTGCC 240
284 GTCGCGAGAGAGACGATCGAACTCCAGACCGCTTCTATGCTCTCTGATTAACCTGCC 343
241 GCTTGAAGGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
344 GTTTTGAAGGTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 403
301 ATCAAGATATCTGCTGATGAGAAAGATTCAACAGATGATCAAGGCTTTAAAGTAAAG 360
404 ATTAAGATATCGGATGAGCAAGAAATTTCACTCAGATGATTAAGGCTTCAAAAGTAAAG 463
361 CACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
464 CATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 522
421 CTCTAAGAAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
523 --TTCTGGAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 580
481 GGGATCCGATGCTTATCGGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 540
581 GGGATCCGATGCTTATCGGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 640
541 ACAAGTGGTTCATTAACACCAAGATGATGATGATGATGATGATGATGATGATGATGATG 600
641 ACAAGTGGTTCATTAACACCAAGATGATGATGATGATGATGATGATGATGATGATGATG 700
601 GATGCAAGGTCATGTTGTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 660
701 GATGTCGGTCATGTTGTTCCAGAGATGATGATGATGATGATGATGATGATGATGATGATG 760
661 GATCCAGGTTCACTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
761 GATCCAGGTTCACTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 820
721 GTCGAAGAATCTCTCCGCTGCTGTCGAAGAGCGGTTGTTGATCTGATGATGATGATGATG 780
821 GTCGAAGAATCTCTCTCCGCTGCTGTCGAAGAGCGGTTGTTGATCTGATGATGATGATG 880
781 CCAATCCGATCATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
881 CCAATCCGATCATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 940
841 GGTGAGGATATACCGAAGAGCGGTTCCCTAAATATTCATTAACCTTAACGACTGCA 900
941 GGTGAGGATATACCGAAGAGCGGTTCCCTAAATATTCATTAACCTTAACGACTGCA 1000
901 AGAAACCCACTTGAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
1001 AGAAACCCCGTTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1060

Qy 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db 1061 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1120
Qy 1021 CAGATCATATCCATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Db 1121 CAGATCATATCCATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1180
Qy 1081 GACTGCGAGAGGCTTGGCATGA 1104
Db 1181 GATTGCGAAGAGCTTTACCTGA 1204

RESULT 5
ATH7312 1457 bp mRNA linear PLN 03-MAR-2000
LOCUS
DEFINITION
Arabidopsis thaliana mRNA for pyruvate dehydrogenase kinase.
ACCESSION
AJ007312
VERSION
AJ007312.1 GI:3641833
KEYWORDS
pdh gene; pyruvate dehydrogenase kinase.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
1 Zou J., Qi O., Karavac V., Marillia E.F. and Taylor D.C.
Effects of antisense repression of an Arabidopsis thaliana pyruvate
dehydrogenase kinase cDNA on plant development
Plant Mol. Biol. 41 (6), 837-849 (1999)
MEDLINE
20199409
PUBMED
10737148
REFERENCE
2 (bases 1 to 1457)

AUTHORS
Zou J.
TITLES
Direct Submision
Submitted (18-JUN-1998) Zou J., Seed Oil Modification Group, Plant
Biotechnology Institute, NRC Canada, S7H 0W9, Saskatoon, CANADA
JOURNAL
Location/Qualifiers
1. .1457

FEATURES
source
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultiivar="Columbia"
/db_xref="taxon:3702"
1. .1457
/gene="pdh"
104..1204
/gene="pdh"
/EC_number="2.7.1.99"
/function="phosphorylates mPDC"
/evidence=experimental
/product="pyruvate dehydrogenase kinase"
/protein_id="CA07447.1"
/db_xref="GI:3641834"
/db_xref="GOA:O82657"
/translation="MAVKKACMPKSLIEDVHKMGKMGKOTGVSILRYMREPKPTER
NLLISQPLAKELPIRVARALIELOTLPGLGSDKRAVLITRMVYLSPFDMRSPRIK
DSGEDKFTQMIKAVVRHNVPMAALVNQKMGKSGNDELHOFIDRPLSLIG
IMLIGQVHLEHNPENPLATVGIHTKMSPMKVFVARSADASI CEREGSAEINLY
GDSFETFPVPHLDMLELVKNSLRAVERFVSDRVAPIRIIVAGIIBVITIKV
SDGGIARSGLPIFTLYSTARNLESDVDLGIADVGTMGVGYGLPIISRLVARY
FGDGIQIISMEGYTDVAILHSLRSDSGRLP"

CDs
gene
CDS
104..1204
/gene="pdh"
/EC_number="2.7.1.99"
/function="phosphorylates mPDC"
/evidence=experimental
/product="pyruvate dehydrogenase kinase"
/protein_id="CA07447.1"
/db_xref="GI:3641834"
/db_xref="GOA:O82657"
/translation="MAVKKACMPKSLIEDVHKMGKMGKOTGVSILRYMREPKPTER
NLLISQPLAKELPIRVARALIELOTLPGLGSDKRAVLITRMVYLSPFDMRSPRIK
DSGEDKFTQMIKAVVRHNVPMAALVNQKMGKSGNDELHOFIDRPLSLIG
IMLIGQVHLEHNPENPLATVGIHTKMSPMKVFVARSADASI CEREGSAEINLY
GDSFETFPVPHLDMLELVKNSLRAVERFVSDRVAPIRIIVAGIIBVITIKV
SDGGIARSGLPIFTLYSTARNLESDVDLGIADVGTMGVGYGLPIISRLVARY
FGDGIQIISMEGYTDVAILHSLRSDSGRLP"

Query Match 79.1%; Score 873.6; DB 8; Length 1457;
Best Local Similarity 87.6%; Pred. No. 3.7e-247;
Matches 967; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

1 ATGGCGGTGAAGAGGCTAGCCGAGATGTTTCCGAAGCTTGAATCCAGAGCTTCACAGA 60
104 ATGGCAGTAAAGAAACCTCGGAAATGTTCCGAAGGTTGATCGAAGATGTTCAAAA 163

```

OY 61 TGGGATGATGATGAGCAGACGGGCGGTGAGCCCTCAGCTACATGATGAGATTCGGTCCACT 120
DB 164 TGGGATGATGATGAGCAGACGGGCGGTGAGCCCTCAGCTACATGATGAGATTCGGTCCAA 223
OY 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGGAGTTCTTTCACAGAGACTTCCGATTCGG 180
DB 224 CCACTGAGAGAAACCTTCTGATCTCGGCGGAGTTCTTTCACAGAGACTTCCGATTCGG 283
OY 181 ATCCGAGAGCGGTGAGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTGCC 240
DB 284 GTCCGAGAGAGCGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTGCC 343
OY 241 GTCTGAGAGATGAGATTCGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTGCC 300
DB 344 GTTTGAGAGATGAGATTCGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTGCC 403
OY 301 ATCAAGATTCCTGATGAGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTGCC 360
DB 404 ATTAAGATTCGATGAGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTGCC 463
OY 361 CACAAACAAGTGTCCATGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTGCC 420
DB 464 CATACATGATGATTCGATGAGATTCGAGACCGTCTGATTCGCTCTGAGAAACCTGCC 522
OY 421 CTGACGAGAAAGCTGATGAGATTCGATGAGATTCGAGACCGTCTGATTCGCTCTGATG 480
DB 523 --TTTGGAAATCTTGAATGAGATTCATGATTCGATTCGATTCGATTCGATTCGATTCG 580
OY 481 GGGATTCGATTCGATTCGAGACCGTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTC 540
DB 581 GGGATTCGATTCGATTCGAGACCGTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTC 640
OY 541 ACAATGAGATTCATACACACCAAGATTCCTCTGAGAGGAGGAGAGAGAGAGAGAGAGAGAG 600
DB 641 ACAGTGGATTAATACACACCAAGATTCCTCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 700
OY 601 GATGAGAGTGTGATTCGAGAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 660
DB 701 GATGAGTGTGATTCGAGAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 760
OY 661 GATCAGATTCATCTTTCGAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 720
DB 761 GATCAGATTCATCTTTCGAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 820
OY 721 GTCAAGAACTCTCTCGGCTGCTGTCAGAGAGCGGTTGTTGACTCTGATGAGGTTGACCA 780
DB 821 GTCAAGAACTCTCTCGGCTGCTGTCAGAGAGCGGTTGTTGACTCTGATGAGGTTGACCA 880
OY 781 CCAATCCGATTCATGTTGCTGATGAGATTCGAGAGATTCGATTCGATTCGATTCGATTCGATTC 840
DB 881 CCAATCCGATTCATGTTGCTGATGAGATTCGAGAGATTCGATTCGATTCGATTCGATTCGATTC 940
OY 841 GGTGAGATTCATGAGAGAGCGGTTCCCTAAATATTCCTTAACCTTAACAGCACTGCA 900
DB 941 GGTGAGATTCATGAGAGAGCGGTTCCCTAAATATTCCTTAACCTTAACAGCACTGCA 1000
OY 901 AGAAACCCACTTGAAGAGATTCGAGAGCGGTTGTTGACTCTGATGAGGTTGACCTGATGAGCT 960
DB 1001 AGAAACCCCGCTTGAAGAGATTCGAGAGCGGTTGTTGACTCTGATGAGGTTGACCTGATGAGCT 1060
OY 961 GGTATGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1020
DB 1061 GGTATGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1120
OY 1021 CAGATCATTCATGAGAGATTCGAGAGCGGTTGTTGACTCTGATGAGGTTGACCTGATGAGCTG 1080
DB 1121 CAGATCATTCATGAGAGATTCGAGAGCGGTTGTTGACTCTGATGAGGTTGACCTGATGAGCTG 1180
OY 1081 GACTCGAGAGAGCTTTGCCATGA 1104
DB 1181 GATTGCAAGAGCTTTACCTGA 1204

```

```

RESULT 6
AJ630654
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 Duplico, B., Martin, I. and Labrador, E.
A pyruvate dehydrogenase kinase is expressed in Cicer arietinum
etiolated epicotyls
Unpublished
2 (bases 1 to 1537)
Labrador, E.
Direct Submission
Submitted (10-MAR-2004) Labrador E., Dpto. Fisiologia Vegetal,
Univ. Salamanca, Campus Miguel de Unamuno, E-37007, SPAIN
Location/Qualifiers
1.1537
/organism="Cicer arietinum"
/mol_type="mRNA"
/cultivar="Pedrosillano"
/db_xref="taxon:3827"
/clone="CampDX"
/tissue_type="etiolated epicotyls"
/clone_lib="CAN-5"
/dev_stage="5 days old seedlings"
1..1537
/gene="Pdk"
147..1250
/gene="Pdk"
/BC_number="2.7.1.99"
/codon_start=1
/product="pyruvate dehydrogenase kinase"
/protein_id="CAG14980.1"
/db_xref="GI:45720178"
/translation="MAAIEAFSKSLIEVHKMGCLKQGVSLRYMMEGSKPTKNL
ISAOFLKEALIRAIIELETLPLYGSGCAVALKQVVDVSDPRDGRSPKNTN
DEKETEYIKAIKRVHNVVPTMLGVQQLKKGAKRMWNEDEPIHGFLEFYSRI
GIRMLIGVHLEHNPPIYVGTHITMSVEVAKNASDEARSI CUREYGSADPINT
YGDPEFPPVPAHLHLMVFEVLVNSLAAVGRYVSDSKSPRIIIVADLEBVTIK
ISDEGGIARSGLPKIFTYLYSTRNPLDEHEDLVADVSTTMAGYGLPISRLYAR
YFGDQLQIISMEGYTDAYLHLSRLGSQEPPLP"

ORIGIN
Query Match 57.9%; Score 639.2; DB 8; Length 1537;
Best Local Similarity 74.2%; Pred. No. 11e-177;
Matches 825; Conservative 0; Mismatches 278; Indels 9; Gaps 1;

OY 2 TGGCGTGAAGAAGGCTGAGAGATGTTTGAAGAGCTTGATGAGAGAGCTTCAAGAT 61
DB 139 TGTGATGATGAGCTGAGAGATGTTTGAAGAGCTTGATGAGAGAGCTTCAAGAT 198
OY 62 GGGAGTGCATGAAGCAGAGCGGCGGTGAGCTCAGGTAATGATGAGAGCTTCCGATTC 121
DB 199 GGGGTTGTTGAACAACCTGCTGAGTCTTAGGTAATGATGAGAGCTTCCGATTC 258
OY 122 CCACTGAGAGAAACCTTCTGATCTCGGCGGAGTTCTTCAAGAGAGCTTCCGATTC 181
DB 259 CTAAGTGAAGAAATTTGATTCGACAGCTTCAAGAGAACTTGCCTATTAAGA 318
OY 182 TCGGAGGCGTGCATGAGATTCGAGAGCGGCTTATGAGCTCTGAGAGAACTGCGG 241
DB 319 TTGCTAGAGAGGCTATGAGCTTGAACCTCTCCCTATGGGTTGTCAAAAAACCTGCTG 378
OY 242 TCTTGAAGGTAAAGATTTGATGAGATTCAGAGAGCTGTTCTTGAAGA 301

```

Db	379	TATTGAAGTTAGGGATTGGATTGTGATGTGATTCATTCCGGAGATTTGATCTTTCCCTGAAA	438
QY	302	TCAGAGATCTGCTGATGATGAAGAGCTTCACACAGATGATCAAGGCTGTTAAAGTAAGCC	361
Db	439	TCAGAAATATAAATGATGATAAAAAGAGTTCACTGAAGATTATCAAGGCCATCAAGTGAAC	498
QY	352	ACAAACAAGTGGTGTCCCATGATGGCTCGGGGTGAACAGCTGAAGAAAGAAATGAAC	421
Db	409	ACAAACAATGGTACTCAATATGGCCTTAGGGTTCAGAGATGTGAAAAAGGCTTGAAC	558
QY	422	-----TCACGAAAAAGCTTGATGAGATTCATCAGTTTCTTGATGGCTTACTACTGT	472
Db	559	CAAAATATGGTTAATGAAGATTTTGTGTGAGATTCATCAGTGTCTTGATGCTTTTACTTGT	618
QY	473	CTCGTATAGGGATCCGATGCTTATTCGGGACAGCATGTTGAGTTGCATATCAACCAAC	532
Db	619	CAMGAATTTGAATTTGGCATGCTTTATAGGGCAGCATGTTGAGTTGCACATCCATCCTC	678
QY	533	CACCTCACACAGTGGGTTACATACACCAAGATGCTCTCTATGAGAGTGGCAAGAGATG	592
Db	679	CACCTTATGTTGTGGGTTATATATACATACAAAATATCTCTCTGTCAGATGGCGAAGATG	738
QY	593	CTAGTGAAGATGCAAGATGCAATTTGTTTCAGAGATGATGTTCTGCTCCGAGATTAACA	652
Db	739	CCAGTGAAGATGCAAGCTCTATATGTCCTTCGATATATGAAGTCTCTGATATCATATA	798
QY	653	TATATGGCGATCCAAAGTTCACTTTCCGATGTTCCGACCCATTTGCATCTTATGTGT	712
Db	799	TTTATGATGATCTCACTTCACTTTTCCGATGTTCCAGATCATTTGCATCTTATGTGTAT	858
QY	713	ATGATTTAGTCAAGAACTCTCCGCTGCTGTCGACAGACCGTTTGTGACTCTGATAGGG	772
Db	859	TTGAGTTGGTTAAGAACTCATTTGCGTCGTCGCTCCAGAACGTTATATGAGATTCGACAAAG	918
QY	773	TTGCACACCAATCCGATCATTTGTCGATGATGAATGAAGATGTTCAATTAAGGTC	832
Db	919	TTTCACCTCCCATTAAGATTTATAGTGTGATGACTGAGAGATTTTCAATTAAGATAT	978
QY	833	CAGATGAAGTGAAGGTATACCGAGAACCGGTCTCCCTTAATAATTCATTAACCTTACA	892
Db	979	CAGATGAAGGAGGTGAATATGACAAAGATGGTCTGCCCAAAATTTTACATATCTGTACA	1038
QY	893	GCACTGCAGAAACCCACTTGAAGAAATGTGAGACTTTGGGAACCGCTATGTCCCCCTGA	952
Db	1039	GTACTGCAGAAACCCATTTGATGAGCATGAAGATCTTGGAGTACTCATGTAGTGTGACAA	1098
QY	953	CTATGGCTGGTATATGATGATGATGCTGCCATATAGTCGCTGTGATGCTGCATATTTGGTG	1012
Db	1099	CTATGGCTGGATATGATGATATGGCTCTCTATATAGTCGTTTATACGCTCGATTTTCGGAG	1158
QY	1013	GAGATTTGCAGATCATATCATATGGAAGATACGGGACGTAGATGCTTACTTGCATCTGTCTC	1072
Db	1159	GATGATCTTCAATATATCTCTATGGAAGATATCGGAACGATGACATCTCCATTTGTGTTC	1218
QY	1073	GTCCTTGAAGACTCGCAGAGAGCTTTGCCATGA	1104
Db	1219	GTTTGGGAGATTCACMAGAACCTTATACCTGTA	1250
RESULT 7			
LOCUS	AY026039	1500 bp	mRNA linear PLN 14-FEB-2001
DEFINITION	Oryza sativa subsp. indica pyruvate dehydrogenase kinase 1 mRNA, complete cds.		
ACCESSION	AY026039		
VERSION	AY026039.1 GI:12829951		
KEYWORDS	Oryza sativa (indica cultivar-group)		
SOURCE	Oryza sativa (indica cultivar-group)		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.		
REFERENCE	1 (bases 1 to 1500)		

FEATURES	source
<p>AUTHORS Yau, C. P., Zhuang, C. X., Yip, W. K. and Zee, S. Y.</p> <p>TITLE Molecular characterization of a cDNA encoding putative pyruvate dehydrogenase kinase in rice</p> <p>JOURNAL Unpublished</p> <p>REFERENCE 2 (bases 1 to 1500)</p> <p>AUTHORS Yau, C. P., Zhuang, C. X., Yip, W. K. and Zee, S. Y.</p> <p>TITLE Direct Submission</p> <p>JOURNAL Submitted (12-DEC-2000) Department of Botany, The University of Hong Kong, Room 75/06, Kadoorie Biological Science Building, Hong Kong, Hong Kong</p>	<p>location/Qualifiers</p> <p>1..1500</p> <p>/organism="Oryza sativa (indica cultivar-group)"</p> <p>/mol_type="mRNA"</p> <p>/cultivar="IR36"</p> <p>/sub_species="Indica"</p> <p>/db_xref="taxon:39946"</p> <p>/dev_stage="first and second meiosis"</p> <p>55..1146</p> <p>/note="OS-PDK1"</p> <p>/codon_start=1</p> <p>/product="Pyruvate dehydrogenase kinase 1"</p> <p>/protein_id="AAK01947.1"</p> <p>/db_xref="GI:12829952"</p> <p>/translation="MASEPVARAAVEEVGWSNKQTVLSRYMFEFSRPTERNLLLSAOLKEPLRIARALESLSPFGSRPAIKVDMVLSRDIARYPEVNRNDILAFNLIKMIKYRNHNVVPMALGNQOLKNEOVRKILTAPEIHEDLPFMSIIGIRLTIQHVLAHDDPPEPGVIGLINTVELNKSIRAAQVQVPMNSDKVPPRIIVADQAEVITIKYGDFTFFPVVSHLMLPELVKNSIRAAQVQVPMNSDKVPPRIIVADQAEVITIKVSDGGGIPRSGPRLFTYLYSTAKNPMDCEBGTVMAGYGLPISRLYARYFEGGDLQIISMEGYTDAYLHLSRLGDSSEPLP"</p>

QY 532 CCACTTACACAGTGGGTTACATACACCAAGATGTCCTATGAGAGTGGCAAGAT 591
 DB 586 GAGCTGGGCTATAGGCTCATTAATACAGAAATGTCCTCATACAGATGCTCAAGCT 645
 QY 552 GCTAGTGAAGATGCAAGTGCATTTGTTTCAGAGATATGTTCTGCTCGAGATTAAC 651
 DB 646 GCCATGGAAGATGCCCGTTCATTTGTTTGAAGGAATATGATCAAGCTCGAGATTAAC 705
 QY 652 ATATATGCGATCCCAAGTTTCACTTTCCGATATGTTCCGATCCCTTTCATCTATGCTG 711
 DB 706 ATCTATGAGACCCCACTTTTACATTTTCATATGTTTATACATCAATCAATCTATGCTA 765
 QY 712 TATGAGTTAGTCAAGAACTCTCTCCGTCCTGTCACAGAGCCGTTTGTGACTGATAGG 771
 DB 766 TTTGATTTGTGAGAACTCTCTGCTGCTGACAGAGAAAGATATATGAAATTCGATTA 825
 QY 772 GTTGACACCAATCCGATATGTTTCTGATGATGATGATGATGATGATGATGATGATGAT 831
 DB 826 GATGTTCTCCAGTTAGAAATATGTTGCTGATGAGGAGAGAGATGATGATGATGATGAT 885
 QY 832 TCAGATGAGGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 891
 DB 886 AGTATGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 945
 QY 892 AGCATGAGAGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 951
 DB 946 AGCATGAGAGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 993
 QY 952 ACTATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1011
 DB 994 ACTATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1053
 QY 1012 GGAGATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1071
 DB 1054 GATGATGAGAGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1113
 QY 1072 CGTGTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1104
 DB 1114 CGGTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1146
 RESULT 8
 AK069157
 LOCUS
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023007C01, full
 accession AK069157
 VERSION AK069157.1 GI:32979181
 KEYWORDS FLI CDNA; CAP trapper.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Eriatridae; Oryzae; Oryza.
 REFERENCE
 AUTHORS
 1
 The Rice Full-length cDNA Consortium, National Institute of
 Agrobiological Sciences Rice Full-length cDNA Project Team;
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
 Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
 Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
 Science Genome Sequencing & Analysis Group; Ohtomo, Y., Murakami, K.,
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
 Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
 Narikawa, R., Sugiyama, A., Mizuno, K., Kobayashi, M., Xie, Q., Lu, M.,
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Mura, J.,
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arahata, T., Fukuda, S.,
 Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
 Kagawa, I., Kondo, S., Komono, H., Miyazaki, A., Otsu, N., Ota, Y.,
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,
 Yoshino, M., and Hayashizaki, Y.
 Collection, mapping, and annotation of over 28,000 cDNA clones from

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

japonica rice
 Science 301 (5631), 376-379 (2003)
 22752273
 12869764
 2 (bases 1 to 1422)

TITLE
JOURNAL

COMMENT

Direct Submission
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
 Agrobiological Sciences, Department of Molecular Genetics, Head of
 Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
 305-8602, Japan (E-mail:skikuchi@ias.affrc.go.jp,
 Tel:81-29-838-7007, Fax:81-29-838-7007)
 This clone is one of the 28K full-length cDNA clones from japonica
 rice. URL: <http://cdna01.dna.affrc.go.jp/cDNA/>
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T.,
 Yamamoto, M.,
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
 Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Mura, J.,
 Mizuno, K., Narikawa, R., Nishikura, J., Oka, M., Ryu, R., Sugano, S.,
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
 Yoshimura, A., Matsubara, K., and Murakami, K.
 Genome Exploration Research Group in Riken Genomic Sciences Center
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
 Akimura, T., Arahata, T., Carninci, P., Fukuda, S., Hanagaki, T.,
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
 Hiraoka, T., Hori, F., Iida, J., Imamura, N., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Komono, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishii, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N.,
 Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
 Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagami, T., Tagawa, A., Takahashi, F.,
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K.,
 Yasunishi, A., and Hayashizaki, Y.
 Location/Qualifiers

FEATURES
source

1..1422
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultiyar="Nidpobare"
 /db_xref="taxon:39947"
 /clone="U023007C01"

ORIGIN

Query Match 49.1%; Score 542.2; DB 8; Length 1422;
 Best Local Similarity 70.1%; Pred. No. 5,9e-149;
 Matches 768; Conservative 0; Mismatches 303; Indels 24; Gaps 2;

QY 840 AGGTGAGGTATTCAGAGAGCGGTCTCCCTAAATATTCATTACCTTACAGACACTGC 899
 |||||
 DB 917 AGGTGCGGGATACCAAGAGCGGCTCCCGAATTTTCACTATCTGACAGACCGC 976
 |||||
 QY 900 AAGAAACCCACTTGAAGAGATGTGACTTGGGAACCGCTGATGTTCCCTGACTATGCC 959
 |||||
 DB 977 AAGAAATCTCTCGACTAGAT-----AGACCTAATACGTAAGGGTAACTATAGGC 1027
 |||||
 QY 960 TGGTATGTTATGCTGTCTGCTTATAGTCCCTGTGATGCTGCTATTTTGTGGAATTT 1019
 |||||
 DB 1028 TGGGATGTTTTCGGGCTCCCGATCAGCCGACTCTGCTGCTATTTTGGCGGCACTT 1087
 |||||
 QY 1020 GCAGATCATATCCATGAAGAGATGAGGAGTGTATGCTTACTTGAATCTGCTGCTGCTG 1079
 |||||
 DB 1088 GCMAATCATCTGCATGAGAGATGAGATGCTTACTTCACTTGTGCGGCTGAG 1147
 |||||
 QY 1080 AGACTGCAAGAGCCTTTGCCATGA 1104
 |||||
 DB 1148 AGATTGAGAGAGCCCTGCTTGA 1172
 |||||

RESULT 11
 AF038585 1380 bp mRNA linear PLN 15-OCT-1998
 LOCUS Zea mays pyruvate dehydrogenase kinase isoform 1 mRNA, nuclear gene
 DEFINITION encoding mitochondrial protein, complete cds.
 AF038585
 VERSION AF038585.1 GI:3746430
 KEYWORDS

SOURCE Zea mays
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
 AUTHORS Thelen, J.J., Muszynski, M.G., Mierzyk, J.A. and Randall, D.D.
 TITLE Molecular analysis of two pyruvate dehydrogenase kinases from maize
 JOURNAL J. Biol. Chem. 273 (41), 26618-26623 (1998)
 MEDLINE 98434573
 PUBMED 9786901

REFERENCE
 AUTHORS Thelen, J.J., Mierzyk, J.A. and Randall, D.D.
 TITLE Direct Submission
 JOURNAL Submitted (15-DEC-1997) Biochemistry, University of Missouri, 117
 Schweitzer Hall, Columbia, MO 65211, USA
 3 (bases 1 to 1380)

REFERENCE
 AUTHORS Thelen, J.J., Mierzyk, J.A. and Randall, D.D.
 TITLE Direct Submission
 JOURNAL Submitted (09-OCT-1998) Biochemistry, University of Missouri, 117
 Schweitzer Hall, Columbia, MO 65211, USA
 Nucleotide sequence updated by submitter
 On Oct 15, 1998 this sequence version replaced gi:3695002.

REMARK
 COMMENT Location/Qualifiers
 FEATURES
 source

1. 1380
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="taxon:4577"
 SS..1146
 /EC_number="2.7.1.99"
 /function="catalyzes the phosphorylation of pyruvate
 dehydrogenase"
 /note="PDK1"
 /codon_start=1
 /product="pyruvate dehydrogenase kinase isoform 1"
 /protein_id="AAC63961.1"
 /db_xref="GI:3746431"
 /translation="MASEPVAVAEVAVRWGAMRQGVSLRYMEFGARPTERTLLT
 ALOPLHEKPIRIARALDLSLPGLSTKPAIKYDMVYSFREISFEPRNOKD
 BLAFTOMKIMIRVHNWVPAIALGVQOLKQIGKAPPGHEIHQIDLRPYMRI
 GTRMLGQVHALHDPDPPEGVGLINTKSPPTVARIASDRAIKRMREYSSPDVI
 YDPPGFTFYVPHLMLFELVKNSLRAVQERYMSDKLAPVRIIVADGADVLIK
 ISDEGGIIPRSGLSRIFTYLYSTAEINPDLQDHNQVTAAGYGIPIRLVARYFGG

CDS

ORIGIN DLQIISMEGVTDVYLHLRLSDSEERLP"

Query Match 46.9%; Score 517.4; DB 8; Length 1380;
 Best Local Similarity 68.4%; Pred. No. 1,3e-141;
 Matches 757; Conservative 0; Mismatches 326; Indels 24; Gaps 2;

QY 10 AAGAAAGCTAGAGAGATGTTTTCGAAGAGCTTATGAGAGACCTTACACAGATGGAGATGC 69
 |||||
 DB 52 AGATAGGCTTCGAGAGCGGCTGCGCGGCGGAGAGAGAGTGGCGGCTGGGGCGC 111
 |||||
 QY 70 ATGAAACAGACGCGGCGTGAACCTCAGTACATGATGATGATGATGATGATGATGATGATG 129
 |||||
 DB 112 ATGCGCAGACCGGGGTGACGCTGCGCTACATGATGATGATGATGATGATGATGATGATG 171
 |||||
 QY 130 AAGAACTTCTGATCTCGGCGCAGTTTCTTCAACAGAGCTTCCATTTGGATCGGAGG 189
 |||||
 DB 172 CGGACCTCTCTCTCGCGCGGAGTTCGCAACAGAGCTCCCATTCGCGATCGGCGCA 231
 |||||
 QY 190 CGTGCATCGAACTCGAGAGCGTGCCTTATGAGGCTCTTGAAGAACTGCGCTTGAAG 249
 |||||
 DB 232 CGGCGCTGACCTCGATCGCTGCTCCCTTCCGCTTCCACCAAGCCGCTATCTCAAG 291
 |||||
 QY 250 GTAAGAGATTGT 309
 |||||
 DB 292 GTGAAGATTGT 351
 |||||
 QY 310 ACTGCTGATGAGAAAGAGTTCACACAGATGATGATGATGATGATGATGATGATGATGATG 369
 |||||
 DB 352 CAGAAAGATGAGCTGCTTCAACCCAGATGATGATGATGATGATGATGATGATGATGATGATG 411
 |||||
 QY 370 GTGTTCCCATGATGCTGCGGTGGAACCAAGCTGAAGAAAG-----GATG 417
 |||||
 DB 412 GTGTTCCCATGATGCTGCGGTGGAACCAAGCTGAAGAAAG-----GATG 471
 |||||
 QY 418 AAATCTACGAAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 477
 |||||
 DB 472 GCGTCCCCCGGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 531
 |||||
 QY 478 ATGAGATCCGATGCTTATCGGCGAGCATGTTGATGATGATGATGATGATGATGATGATGATG 537
 |||||
 DB 532 ATGAGATCCGATGCTTATCGGCGAGCATGTTGATGATGATGATGATGATGATGATGATGATG 591
 |||||
 QY 538 CACACAGTGGTTACATACACACCAAGATGCTCTGATGATGATGATGATGATGATGATGATGATG 597
 |||||
 DB 592 GGTGATAGGGCTTATTAACACAAATATGTACCCATGATGATGATGATGATGATGATGATGATG 651
 |||||
 QY 598 GAAGATGCAAGTGCATTTGTTTTCAGAGATGATGATGATGATGATGATGATGATGATGATGATG 657
 |||||
 DB 652 GAAGATGCAAGTGCATTTGTTTTCAGAGATGATGATGATGATGATGATGATGATGATGATGATG 711
 |||||
 QY 658 GCGATCCAAATTTTCACTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 717
 |||||
 DB 712 GGTATCCGGGTTTATCATTTTCCATATGTTTACACCAATCTCATGATGATGATGATGATGATG 771
 |||||
 QY 718 TTAGTCAAGAACTCTCTCGGCTGCTGCAAGAGCGGTTTGTGACTCTGATAGGTTGCA 777
 |||||
 DB 772 CTGCTGAAGAACTCTCTCGGCTGCTGCAAGAGCGGTTTGTGACTCTGATAGGTTGCA 831
 |||||
 QY 778 CCACCAATCCGATCATTTGTTTTCAGAGATGATGATGATGATGATGATGATGATGATGATGATG 837
 |||||
 DB 832 CCTCCAGTTCCGATCATTTGTTTTCAGAGATGATGATGATGATGATGATGATGATGATGATGATG 891
 |||||
 QY 838 GAAGTGAAGTATACCGGAAGCGGCTCCCTAAATATTTCACTTACCTTACAGACACT 897
 |||||
 DB 892 GAAGTGAAGTATACCGGAAGCGGCTCCCTAAATATTTCACTTACCTTACAGACACT 951
 |||||
 QY 898 GCAAGAAACCACTTGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 957
 |||||
 DB 952 GCAAGAAACCACTTGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 999
 |||||
 QY 958 GCTGTTATGTTATGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1017
 |||||

RESULT 12	LOCUS	DEFINITION	ACCESSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
QY	AK100033	Oryza sativa (japonica cultivar-group) cDNA clone: J013149J07, full insert sequence.	AK100033	FLI -CDNA; CAP trapper.	ORIYAZ SATIVA (JAPONICA CULTIVAR-GROUP)	Oryza sativa (japonica cultivar-group)	1	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team;
QY	1018	TTGCAGATCAATATCCATGAGAGATACGGAGCTGATGCTTACTTGCCTTGTCTGCTTT	1535 bp					Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Db	1060	TTGCAGATCATCTCTATGAGAGATATGGAATGAGTATGCTTACTTCCACTTTCACGGCTG						Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
QY	1078	GGAGACTCGCAGAGACTCTTGGCCATGA	1104					Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Db	1120	GGAGATTTGGAGAACCTTGGCCGTAA	1146					Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Ohtomo, Y., Murakami, K.,
								Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
								Kurosaki, K., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
								Marikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nikura, J.,
								Ikedu, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
								Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;
								Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
								Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Kawai, J., Kawamata, S.,
								Kagawa, I., Kondo, S., Kono, H., Miyazaki, A., Oshio, N., Ota, Y.,
								Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
								Yoshino, M., and Hayashizaki, Y.
TITLE								Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL								Science 301 (5631), 376-379 (2003)
MEDLINE								27352273
PUBMED								12869764
REFERENCE								2 (bases 1 to 1535)
AUTHORS								Aachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
								Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
								Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiyama, K., Hiraoka, T.,
								Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
								Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
								Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
								Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
								Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Kono, H., Kondo, M.,
								Koya, S., Kurihara, C., Kurosaki, K., Kusumegi, T., Li, C., Lu, M.,
								Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
								Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
								Namiki, T., Narikawa, R., Nikura, J., Nishii, K., Nomura, K.,
								Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
								Osato, N., Ota, Y., Ohtomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
								Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
								Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
								Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, T., Tagami, T.,
								Tagawa, A., Takahashi, F., Takaki-Akita, S., Tanaka, T., Tomaru, A.,
								Toyoda, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
								Yamada, H., Yamamoto, M., Yasuniishi, A., Yazaki, J., Yokomizo, S., and
								Yoshimura, A.
TITLE								Direct Submission
JOURNAL								Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
								Agrobiological Sciences, Department of Molecular Genetics, Head of
								Laboratory of Gene Expression, 2-1-2 Kamondai, Tsukuba, Ibaraki

Query Match	Best Local Similarity	Matches	Conservative	Score	DB	Length
Query Match	46.3%	68.0%	Prod. No. 8e-140	511.4	DB 8	1535
Best Local Similarity	68.0%	Prod. No. 8e-140				
Matches	754	Conservative	0	Mismatches	331	Indels 24; Gaps 2
3	GGCGGTGAAGAAGCTAGCGAGATGTTTTCGAAGAGCTTGAATCGAGAGCTTACAGATG	62				
170	GGGGATGATGCGCCTGACCTCGGAGCGCGGCGGAGGCGCGGTGGGGAGAGAGTGGCCGCGT	229				
63	GGGATGATTAACCAACGCGCGCTGAGCTCAGATGATGATGAGAGTTGGTTCCATCC	122				
230	GGGAGGATGAGGCAACGCGGGGTGACGCTGGCGATACATGATGAGAGTTGGGGCGCGCC	289				
123	CACGTGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAACAGAGAGTTCCGATTCGGAT	182				
290	CACGAGAGGAACCTCTCGTGGCGGTGGCGGAGTTCTCGGAGAGAGTCCCATCCGAT	349				
183	CGGAGCGCTGCGATGAACTGAGACGCTGCTTATGCGCTTCTTGAAGAACCTGCGGT	242				
350	CGCGCCCGCGCGCTGACCTGACCTCTCCCTCCCTCGGCGCTCTCCACCAAGCCGCGCAT	409				
243	CTTGAAGTAAAGATTTGGATGATGAGATCAATTCAGAGGACATGAGAGGTTTCCGAGAT	302				
410	CTTCAAGTGAAGATTTGGATTTGGATCTCATTTCCGCACTTACGATGCTTCCCGAGAGT	469				
303	CAGATATCTGCTGATGAGAAAGATTCACAGATGATCAAGGCTGTTAAAGTAAAGCA	362				
470	GAGGAACCGGAGTATGAGCTTGCATTCACCGAGATGATCAAGATGATTCAGATGCGGCA	529				
363	CAGCAACGTGATTTCCATGATGAGCTCTGCGGTGAAACAGCTGAAGAAAG-----	412				
530	CACCAATGATGAGCCACCATGCGCTTGGGTGGTGGGAGCTGAAGAAAGCTGGGTGG	589				
413	--GATGAACCTCTACGAAGAGCTGATGAGATTCATCAGTTTCTTGAATCGCTTCACTT	470				


```

Db      721 GTTCTCCAGTTAGAAATATATGCTTCTATGG-----GGGAGAACATATCAAGTTTACT 774
Qy      835 GATGAAGGTGAGATATACCGAAGAGCGTCTCCCTAAATATTCATTACCTTACAGC 894
Db      775 GATGAAGGTGAGATATACCGAAGAGCGCTTCCAGAAATTTTCACTACCTATATAGC 834
Qy      895 ACTGCAAGAAACCCACTTGAAGAGATGTGACTTTGGGACCGCTGATGTTCCCTGACT 954
Db      835 ACTGCAAGAAATCCACTT-----GATATGATGTCCTCCTAGTGAAGAGTAACT 882
Qy      955 ATGCTGTGTTATGTTATGTTCTGCTATATAGTCTGCTTATGCTGCTATTTTGTGA 1014
Db      883 ATGCTGTGATATGCTATGTTCTTCCATTCAGTGGCCGTATGCTCCATATTTTGTGT 942
Qy      1015 GATTTGAGATCATATCATGAGAGATACGGGACTGATGCTTACTTCTGCACTTGTCTGT 1074
Db      943 GACCTGCAATCATCTCTATGAAAGATACGGCACCGATGCTTACCTGCACTTGTCAAG 1002
Qy      1075 CTGAGAGACTGCGAGAGCCCTTTCATGA 1104
Db      1003 TTGGAGATTGAGAGAGCCCTTCTTGA 1032

```

```

RESULT 14
LOCUS      ARS66929      899 bp      DNA      linear      PAT 08-OCT-2004
DEFINITION Sequence 965 from patent US 6768041.
ACCESSION  ARS66929
VERSION     ARS66929.1 GI:53983965
KEYWORDS
SOURCE      Unknown.
            Unclassified.
ORGANISM    1 (bases 1 to 899)
            Strubala,T. and Nieuwenhuizen,N.
REFERENCE   1 (bases 1 to 899)
            Compositions isolated from plant cells and their use in the
            modification of plant cell signaling
            Patent: US 6768041-A 965 27-JUL-2004;
            Location/Qualifiers
            1..899
            /organism="unknown"
            /mol_type="genomic DNA"

```

```

ORIGIN
Query Match      29.4%; Score 324.2; DB 6; Length 899;
Best Local Similarity 76.0%; Pred. No. 2.2e-84;
Matches 414; Conservative 0; Mismatches 128; Indels 3; Gaps 1;

```

```

Qy      560 CCAAGATGCTCTCTATGAGAGTGGCAAGATCTGTAAGATGCAAGTGGATTGTT 619
Db      144 CGAAGATGCTCCAGTTGAGGTGCACTATGCTATTTGAGATGCTCGTCTATTGTT 203
Qy      620 TCAGAGATATGTTCTGCTCCGAGATAAACATATATGCGATCCAGATTTCATTTC 679
Db      204 TGGGCGAGATATGCGACTGCTCCGATATTAATATATGCGATCCAGATTTCATTTC 263
Qy      680 CGATATGTCGACCCATTGTCATCTTATGCTGATGATGATGATGATGATGATGAT 739
Db      264 CGATATGTCGACCCATTGTCATCTTATGCTGATGATGATGATGATGATGATGAT 323
Qy      740 CTGTCAGAGAGCGTTGTTGATCTGATAGAGGTTGACACCAATCCGTAATCTTTG 799
Db      324 CTGTCAGAGAGCATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 383
Qy      800 CTGATGAATCGAAGATGTTTACATTAAGATCTCAGATGAGATGAGATGATGAGATG 859
Db      384 CTGATGAAGACGAAGAGCTTACATCAAGTCTCAGATGAGAGGAGGAGTCCAGAGAA 443
Qy      860 GGGGTCTCCCTAAATATTCATCTTACCTTACAGCAGCTGCAAGAAACCCACTTGAAG 919
Db      444 GTGGTCTTCCCAAAATCTTCACTATCTGATATGAGCACTGCGAAAAACCCCTGATGAGA 503

```

```

Qy      920 ATGTGACTTGGAGACCGCTGATGTTCCCTGACTATGCTGTGTTATGTTATGTTCTGC 979
Db      504 ACTCAGATCTTGGAAATAGCTGATAAT--GTGACTATGCGCGGTTATGTTATGGCTTC 560
Qy      980 CTATTAGTGCCTGTTATGCTCCGCTATTTTGTGAGATTTTGCAGATCATATCTGAGAG 1039
Db      561 CAATCAGCGCTGTTATGCTCGTACTTGGTGTGATCCTGCAAAATATTTTCTATGAGAG 620
Qy      1040 GATACGGAGCTGATGCTTACTTGCATCTGCTGCTGTTGAGACTGCGAGAGCCCTTTC 1099
Db      621 GATATGGAAGTATGATGATACCTTCATTATCTGCTCGCAATTCGCAAGAACCGCTTC 680
Qy      1100 CATGA 1104
Db      681 CGTGA 685

```

```

RESULT 15
LOCUS      ATAC011623      86022 bp      DNA      linear      PIN 30-OCT-2002
DEFINITION Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence,
            complete sequence.
ACCESSION  ATAC011623
VERSION     ATAC011623
KEYWORDS    HTG.
SOURCE      Arabidopsis thaliana (thale cress)
            Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
            1 (bases 1 to 86022)
            Lin,X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B.,
            Rensing,C.M., Koo,H., Fujii,C.Y., Uteckack,T.R., Batmsstead,M.E.,
            Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
            Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence
            Unpublished
            2 (bases 1 to 86022)
            Lin,X. and Kaul,S.
            Direct Submission
            Submitted (08-OCT-1999) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org
            3 (bases 1 to 86022)
            Lin,X.
            Direct Submission
            Submitted (24-JAN-2001) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
            On Jan 24, 2001 this sequence version replaced gi:12280803.
            Address all correspondence to:
            Xiaoying Lin
            The Institute for Genomic Research
            9712 Medical Center Dr.
            Rockville, MD 20850, USA
            e-mail: xlin@tigr.org
            BAC clone F24P17 is from Arabidopsis chromosome III and is near the
            molecular marker ma403.
            The orientation of the sequence is from SP6 to T7 end of the BAC
            clone.

```

Genes were identified by a combination of three methods: Gene prediction programs including GPRIL (available by anonymous ftp from arthur.epm.ornl.gov), GeneFinder (Phil Green, University of Washington), Genscan (Chris Burge, <http://www.cbs.dtu.dk/netpene/cbsnetpene.html>), and NetPlantGene (<http://www.cbs.dtu.dk/netpene/cbsnetpene.html>), searches of the complete sequence against a peptide database and the Arabidopsis EST database at TIGR (<http://www.tigr.org/tdb/at/est.html>). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted to be more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy).

http://genome.wustl.edu/eddy/cRNAcan-SE/. Simple repeats are identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Regions of genomic sequence that are not annotated as genes but have predicted exons by GRail are annotated as misc features.

FEATURES

```
source
    1. 86022
       /organism="Arabidopsis thaliana"
       /mol_type="genomic DNA"
       /cultivar="Columbia"
       /db_xref="taxon:3702"
       /chromosome="III"
       /map="m1403"
       /clone="F24P17"
       <283..1822
       /gene="F24P17.1"
       /note="similar to pyruvate dehydrogenase kinase
       GB:AC97601, [Arabidopsis thaliana]"
       join(<283..529,638..818,916..1063,1101..1142,1160..1376,
       1511..1822)
       /gene="F24P17.1"
       join(<283..529,638..818,916..1063,1101..1142,1160..1376,
       1511..1569)
       /gene="F24P17.1"
       /codon_start=1
       /product="putative pyruvate dehydrogenase kinase, 5'
       partial"
       /protein_id="AA08568.1"
       /db_xref="GI:6437541"
       /translation="VRDWYLESFRDRAPEIKDSDGDKFTQMIKAVKVRNNVVPM
       MALGVOLKKMGNSGNDIEHQFLDRFYLSRIGIRMLIGQHVLEHNPPLHTVGYIH
       TKMSPEVVRNAGSDARSICFREYSGAPENINIGDSFPEYPTLHLMAYLVNS
       LVAODRPFVDSQDVARPIRIIVADIGEDYIKFRSLHAFDPIIVVDSGGLIARSG
       LRIFTYVISTANRPLEEDVDGIADVPVTMAGYIGLPSRLIARYFGSDLOIISME
       GGETDAYLHLSRIGDSQELP"
       complement(2022..2051)
       /rpt_family="POLY A"
       /note="exon predicted by xgrail, quality
       excellent_shadowexon"
       2057..2160
       /note="exon predicted by xgrail, quality excellent"
       2287..2363
       /note="exon predicted by xgrail, quality excellent"
       <2630..>6599
       /gene="F24P17.2"
       /note="similar to RNA helicase (DRH1) GB:BAA28347
       [Arabidopsis thaliana]; Pfam HMM hits: WW/TP5/WMP domain
       containing proteins, DEAD and DEAD box helicases,
       Helicases conserved C-terminal domain"
       join(<2630..3712,3813..4007,4256..4324,4406..4675,
       4767..5117,5215..5492,5579..>6599)
       /gene="F24P17.2"
       join(2630..3712,3813..4007,4256..4324,4406..4675,
       4767..5117,5215..5492,5579..6599)
       /gene="F24P17.2"
       /codon_start=1
       /product="putative RNA helicase"
       /protein_id="AA08584.1"
       /db_xref="GI:6437557"
       /translation="MATEEDTPASAGRPVAPEDPTLPQPMKGLIDSGTGLIYYMNET
       INTQYVRPSAPRPHSATTPKLAQIPVPSOQAOHOEOKXPGHVSQOOGFOOQOQ
       PSQHYRPMQOHPAOCMPQOQOQPOOQOQOQPMRPHGHSPVOTYOTPTTQOQOQ
       MONHSCMQOQLSHQYLAHQHMGFRPHMQTQGLONSHOTPCGPHGQCFPQOQYTN
       SLAKREGDFHGKCTGFSQPHLPVNSERSPQNTREAAASQKTNANLAMAQKNG
       PANAANVTOFOQGANLHQQLGPRAPONDQMLKQSVSPQGNNTYENNLQKNG
       GNDSYVMRMEYVVGAGQPLHPAPKXDIRISGSPPTNADPAMGQTHGYGAGAF
       PMSLVRPHVTSQDVPLSPVEIYRKQHVTTGGINIPAPYTFSSGGLPEIIRLEL
       LSAQSPSPRIQKQTPALQSDIVALAKTQSGKTLGYLIPAFILLRHGNDNRSGP
       TVLILAPTRRLAQIODEALRFGSSRSISCTCICXGAKPKQPKELERCADIVATRG
       RLNDILMKRVIDQVSLVLADSRMLDGFEPQIRKTYNELPRRQTLMTATPK
       EVAKIASDLVNPQVNIQGVDELANKAITQYVEVVPQMEKERRLEQILRSQERSK
       VIFCSTKRLCDHLARSVGRHFGAVVIHQDKTGQERDWWLQFRSGKS CVL1ATDVAA
```

```
repeat_region
    5843..5962
       /rpt_family="(TGC)n"
       complement(7077..>8090)
       /gene="F24P17.3"
       /note="similar to hypothetical proteins GB:CAB16818,
       GB:AA055277"
       complement(7077..>8090)
       /gene="F24P17.3"
       complement(7254..8090)
       /note="unknown protein"
       /codon_start=1
       /protein_id="AA08569.1"
       /db_xref="GI:6437542"
       /translation="MASIYSSLTLYLVNHPYISNFWIGETLGSVFPFVVSVLYL
       SATFLRSALDLSPLSPRIKPTLVHSLILCLSLVMAVGCTLITSASDPMAR
       RLHAICPEVDKPNGLPEFMAQVFLSKILFEGDTILLILKRSIQRLSFLVYHAT
       VVMCYLMLRTQSMFPIALVNTSVHVMYGYFLCAQGRKPKRLVTDQIVQFV
       SPFGSLGSMWLREHLFGSGCTGIWGCENAFNALLALPSNFHSHKRYKKPTREDGK
       SP"
       8262..8362
       /note="exon predicted by xgrail, quality
       marginal_shadowexon"
       complement(8375..8450)
       /note="exon predicted by xgrail, quality marginal"
       complement(<8719..>9615)
       /gene="F24P17.4"
       /note="similar to hypothetical proteins GB:CAB16818,
       GB:AA055277"
       complement(<8719..>9615)
       /gene="F24P17.4"
       complement(8719..9615)
       /gene="F24P17.4"
       /note="unknown protein"
       /codon_start=1
       /protein_id="AA08570.1"
       /db_xref="GI:6437543"
       /translation="MASVSTLTLYLVHHPYIANFWTGETLGSVFPFVVSVLYL
       SATFLRTYVDSLPITGPRILKPTLVHSLILFLSLVMAVGCTLITSDDPKALF
       DAVCFPLDVKPKPLPEFMAQVFLSKILFVDLILLKSLTQRLSLVYTHATYVI
       LCYLMLRTQSMFPGVLVNTSVHVMYGYFLCAIGSRPKMKRLVTFQMVQFAGM
       GIGAAWMLPEHFGSGAGIWTYFENGFTASLLALFVNFHSHKRYKTTSPLYKIES
       FFIHGERANAKATLFSKND"
       9798..9853
       /note="exon predicted by xgrail, quality good_shadowexon"
       complement(10146..11457)
       /gene="F24P17.5"
       /note="similar to putative protein GB:CAB45783"
       /pseudo
       complement(12637..12715)
       /note="exon predicted by xgrail, quality
       marginal_shadowexon"
       <12993..>16211
       /gene="F24P17.6"
       /note="similar to putative Anion exchanger family members:
       GB:AA055295 [Arabidopsis thaliana]"
       join(<12993..13102,13175..13257,13357..13549,13667..13805,
       13900..14065,14182..14339,14418..14612,14700..14795,
       14879..15013,15093..15400,15493..15763,15855..>16211)
       join(12993..13302,133175..13257,13357..13549,13667..13805,
       13900..14065,14182..14339,14418..14612,14700..14795,
       14879..15013,15093..15400,15493..15763,15855..16211)
       /gene="F24P17.6"
       /note="unknown protein"
       /codon_start=1
```


Query Match 29.3%; Score 324; DB 8; Length 86022;
 Best Local Similarity 64.5%; Pred. No. 5.4e-84;
 Matches 725; Conservative 0; Mismatches 95; Indels 304; Gaps 4;

```

QY 224 TCTCTGAGAAACCTCCCGCTTTGAAAGGTAAGAGATTGGTATGTGAGTCATTCAGGACA 283
DB 257 TGTCTCAATTTTCATCTTGTATTTGTAGTGGCGGATTTGGTATTTGGAACTTTCAAGGACA 316
QY 284 TGAGAGCGTTTCTGAGATCAAGATTAAGTATGAGTGAAGAGTTCACAGATGATCA 343
DB 317 TGAGAGCATTTCTGAGATTAAGATTTGGGTGACGAGAAAGATTTCACTAGATGATTA 376
QY 344 AGCGTTTAAAGTAAGGACAAACACGTTTCCCATGATGCTCTGGGTGTAACGAGC 403
DB 377 AGCGTCAAAAGTAAGCATTAACATGTGTTCCCATGATGCTTTGGGTGTAATCAGC 436
QY 404 TGAAGAAAGGAATGAATACTACGAAAGCTTGAATGATTCATCACTTTCTTGATGCT 463
DB 437 TCAAGAAAGGAATGAA--TTCTGAAATCTTGATGAGATTCATCAGTTCTTGATGCTT 493
QY 464 TCTACTTGTCTGATAGGATCCGTAAGCTTAT----- 497
DB 494 TCTACTTGTGCGAATCGGATCCGATGCTTATTTGTAAGACTTGTTCATTTTTTTT 553
QY 498 ----- 497
DB 554 TCAATCTCTTTTCCCTGACTGAAAGTTTGTGCTGATGTGTTGTTAATACTTG 613
QY 498 -----CGGCGACATGTTGAGTTGATTCATTAATCCAAACCCACAC 535
DB 614 TTGCGTATCATGAAATTTCTCAGGGGACACGTTGAGTTGATTCATTAATCCAAATCCACCGC 673
QY 536 TTCAACAGTGGGTTACATACACACCAAGATGTCTCTATGAGTGGCAAGAGATGCTA 595
DB 674 TTCAATACAGTGGGTTATATACACCAAGATGTCTCTATGAGTGGCAAGAGATGCTA 733
QY 596 GTGAAGATGCAAGTCAATTTGTTTCAGAGATGTGTTCTGCTCCGAGATTAACATAT 655
DB 734 GTGAAGATGCTCGTCAATTTGTTTCGAGATGACGTTCTGCAACCGAATAACATAT 793
QY 656 ATGGCGATCCAAAGTTTCACTTTTCC----- 680
DB 794 ATGGCGATCCCAAGTTTCACTTTTCCCGAGTTTCAAGTTGACTGCAATTTTATTTCTTA 853
QY 681 ----- 680
DB 854 GATTCAGTTTGTAGCTTTGTTCTCAATGTGTTACCAATATCTTTTCTTCTACTCTGT 913
QY 681 --GTATGTTCCGACCATTTGCATCTTATGTGTATGATTAAGTCAAGAACTCTTCGCT 738
DB 914 AGGTATGTTCCAAAGCATTTGCACTTATGATGTATGAGTCAAGAACTCTCTACGCT 973
QY 739 GCTGTCCAAAGCGGTTTGTGACTGTATGAGGTTGCAACCAATCCGTAATCATTTT 798
DB 974 GCTGTCCAAAGCGGATTTGTGACTGTATGAGGTTGCAACCAATCCGCAATATATGTT 1033
QY 799 GCTGATGGAATCGAAGATTTACATTA----- 826
DB 1034 GCTGATGGAATCGAAGATTTACTATTAAGTTGTTACCTGCACTTTTTCCTCTT 1093
QY 827 ----- 826
DB 1094 TTGCTAGCCATTTCTTTTGTGATCGGTTTGAATCCAAATCATTTGTGAGTTAATA 1153
QY 827 ----AGTCTCAGATGAGGTGAGTATACGAGAAAGCGTCTCCCTAAATATTTCACT 882
DB 1154 TAACAGGTCTCAGATGAGGTGAGTATAGCAAGAAAGCGTCTTCCGAATATTTCAAC 1213
QY 883 TACCTCTAAGCACTGCAAGAAACCACTTGAAGAGATGTGACTTGGAAACGCTGAT 942
DB 1214 TATCTTTACAGCACTGCAAGAAACCGCTTGAAGAGATGTGATTTAGAAATAGCTGAT 1273

```

```

QY 943 GTTCCCTGACTATGCTGGTTATGTTATGTTGCTGCTTATTTAGTCTGTTATGCTCC 1002
DB 1274 GTTCCGTTGACTATGCTGGATATGTTATGCTTTCATTAATGTTGCTGTTATGCTCGA 1333
QY 1003 TATTTGGTGAATTTGCAATCATATTCATGAGAGATTAACG 1046
DB 1334 TATTTGGTGAATTTGCAATCATATTCATGAGAGATTAACG 1377

```

Search completed: April 12, 2005, 06:53:52
 Job time : 5153 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2005, 19:30:52 ; Search time 698 Seconds

(without alignments)
9363.029 Million cell updates/sec

Title: US-10-642-531-1

Perfect score: 1104
Sequence: 1 atcgcggtgagagagctag.....cgcagagaccttgcacatga 1104

Scoring table: IDENTITY_NIC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: geneseqn19808:*
2: geneseqn19908:*
3: geneseqn20008:*
4: geneseqn20018:*
5: geneseqn20018:*
6: geneseqn20028:*
7: geneseqn20028:*
8: geneseqn20038:*
9: geneseqn20038:*
10: geneseqn20038:*
11: geneseqn20038:*
12: geneseqn20048:*
13: geneseqn20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1104	100.0	1104	12	ADL71332 Brassica
2	1095.6	99.2	1104	12	ADL71333 Brassica
3	1062.4	96.2	1104	12	ADL71335 Brassica
4	1062	96.2	1104	12	ADL71334 Brassica
5	873.6	79.1	1457	2	AAV49818 A. thaliana
6	524.6	47.5	1602	5	AAD14380 Maize pyr
7	428.2	40.6	1332	5	AAD14379 Maize pyr
8	328.2	29.7	591	13	ACNS4458 Cotton an
9	244.2	22.1	584	13	ADM80480 Maize CDN
10	236.2	21.4	464	3	AAC36807 Arabidops
11	208.2	18.9	371	3	AAV9663 Eucalyptu
12	177.8	16.1	1523	13	ADG49235 Bacterial
13	167.6	15.2	1407	8	ABT21104 Aspergill
14	167.6	15.2	1410	8	ABT19284 Aspergill
15	151.8	13.8	1625	8	ABT20506 Aspergill
16	151.8	13.8	1625	8	ABT18690 Aspergill
17	151.8	13.8	1625	8	ABT19910 Aspergill
18	151.8	13.8	1625	8	ABT18096 Aspergill
19	149.8	13.6	474	13	ACNS3319 Cotton an
20	121.4	11.0	236	10	ADB76050 Tomato pl

21	120.2	10.9	1536	6	ABZ32464
22	111.4	10.1	695	13	ADG60834
23	110.2	10.0	285	6	ABT76427
24	108.4	9.8	1476	10	ADB69820
25	80.8	7.3	768	12	ADT40939
26	80.8	7.3	1268	5	AA880808
27	80.8	7.3	1470	12	ADL61198
28	80.8	7.3	1599	13	ADQ88173
29	80.8	7.3	1599	13	ADG82586
30	79	7.2	1507	10	ADB53519
31	79	7.2	1829	13	ACN39976
32	79	7.2	1866	2	AAV06836
33	71.8	6.5	1798	10	ACA56530
34	71.8	6.5	1798	11	ADP65448
35	71.8	6.5	1798	12	ADP56326
36	71.8	6.5	1798	13	ADG92844
37	71.8	6.5	4889	6	AA894881
38	71.8	6.5	4889	9	ACH03818
39	70	6.3	466	3	AAV15579
40	66.8	6.1	595	10	ADD33813
41	64.8	5.9	2020	10	ADB69459
42	64.8	5.9	4020	10	ADB69098
43	63.6	5.8	1146	12	ADK71915
44	62.6	5.7	1435	10	ADB58761
45	62.6	5.7	1435	10	ADB53453

ALIGNMENTS

RESULT 1
ADL71332
ID ADL71332 standard; cDNA, 1104 BP.

XX ADL71332;

XX 20-MAY-2004 (first entry)

XX Brassica napus pyruvate dehydrogenase kinase (PDHK) cDNA.

XX Pyruvate dehydrogenase kinase; PDHK; fatty acid; oil content; flowering;

XX vegetative growth; respiration; transgenic plant; gene therapy;

XX plant protectant; growth regulator; gene; ss.

XX Brassica napus.

XX Key Location/Qualifiers

XX CDS 1..1104

XX FT /*tag= a

XX FT /product= "PDHK protein"

XX US2004033606-A1.

XX 19-FEB-2004.

XX 16-AUG-2002; 2002US-00222075.

XX 16-AUG-2002; 2002US-00222075.

XX (ZOU/) ZOU J.

XX (TAYLOR) TAYLOR D C.

XX (MARI) MARI LIA E.

XX ZOU J, Taylor DC, Marillia E;

XX WPI; 2004-246636/23.

XX P-PSDB; ADL71336.

XX New purified or recombinant nucleic acid encoding a Brassica pyruvate

XX dehydrogenase kinase protein, for manipulating pyruvate dehydrogenase

XX activity in plants; e.g. seeds with increased or decreased fatty acid and

XX oil content.

PS Claim 2; SEQ ID NO 1; 14bp; English.

CC The present invention relates to novel *Brassica pyruvate dehydrogenase*
 CC kinase (PDHK) proteins and polynucleotides encoding such proteins.
 CC Sequences of the invention are useful for manipulating pH activity in
 CC plants resulting in seeds with increased or decreased fatty acid and oil
 CC content, plants exhibiting early or delayed flowering time, plants with
 CC increased or decreased vegetative growth, plants with root systems better
 CC to withstand low soil temperatures or frost, plants with tissues
 CC exhibiting higher or lower rates of respiration or plants exhibiting an
 CC enhanced capacity to accumulate biopolymers which rely on acetyl groups
 CC as precursors. They are also useful for generating transgenic plants and
 CC in gene therapy. The present sequence is *Brassica napus pyruvate*
 CC dehydrogenase kinase (PDHK) cDNA.

CC Sequence 1104 BP; 281 A; 231 C; 294 G; 298 T; 0 U; 0 Other;

Query Match 100.0%; Score 1104; DB 12; Length 1104;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGTGAAGAGGCTGAGAGATGTTTTCGAAGAGCTTGATCGAGAGCTTCAAGA 60
 Db 1 ATGGCGGTGAAGAGGCTGAGAGATGTTTTCGAAGAGCTTGATCGAGAGCTTCAAGA 60
 QY 61 TGGGATGATGAG 120
 Db 61 TGGGATGATGAG 120
 QY 121 CCCACTGAG 180
 Db 121 CCCACTGAG 180
 QY 181 ATCGGAG 240
 Db 181 ATCGGAG 240
 QY 241 GTCTTGAAGTGAAG 300
 Db 241 GTCTTGAAGTGAAG 300
 QY 301 ATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 Db 301 ATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 QY 361 CACAAACAGTGTGCTCCATGATGCTCTGGGTGTGAACAGCTGAAGAGAGATGAA 420
 Db 361 CACAAACAGTGTGCTCCATGATGCTCTGGGTGTGAACAGCTGAAGAGAGATGAA 420
 QY 421 CTCTAG 480
 Db 421 CTCTAG 480
 QY 481 GGGATCCGTATGCTTATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 Db 481 GGGATCCGTATGCTTATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 QY 541 ACAGTGGTATACATACACACAAAGATGCTCTATGAGAGGTGCAAGAGATGCTA 600
 Db 541 ACAGTGGTATACATACACACAAAGATGCTCTATGAGAGGTGCAAGAGATGCTA 600
 QY 601 GATGAGAGTGAATGTTTTCAGAGAGTGTGCTCCGAGATTAACATATATGAGC 660
 Db 601 GATGAGAGTGAATGTTTTCAGAGAGTGTGCTCCGAGATTAACATATATGAGC 660
 QY 661 GATCAGAGTTTCACTTTCCGATGTTCCAGACCATTTGAGTATGATGATGATGAT 720
 Db 661 GATCAGAGTTTCACTTTCCGATGTTCCAGACCATTTGAGTATGATGATGATGATGAT 720
 QY 721 GTCAAGAGTCTCTCGGTGCTGTCACAGAGCGGTGTTGACTGTGATAGGGTTGACA 780
 Db 721 GTCAAGAGTCTCTCGGTGCTGTCACAGAGCGGTGTTGACTGTGATAGGGTTGACA 780

QY 781 CCAATCCGATCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 Db 781 CCAATCCGATCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 QY 841 GGTGAGAGTATACCGAG 900
 Db 841 GGTGAGAGTATACCGAG 900
 QY 901 AGAACCACCTTGAAG 960
 Db 901 AGAACCACCTTGAAG 960
 QY 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 Db 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 QY 1021 CAGATCATATCCATGAG 1080
 Db 1021 CAGATCATATCCATGAG 1080
 QY 1081 GACTCGAG 1104
 Db 1081 GACTCGAG 1104

RESULT 2

ADL71333 standard; cDNA; 1104 BP.

ADL71333;

20-MAY-2004 (first entry)

Brassica rapa pyruvate dehydrogenase kinase (PDHK) cDNA.

Pyruvate dehydrogenase kinase; PDHK; fatty acid; oil content; flowering;

vegetative growth; respiration; transgenic plant; gene therapy;

plant protectant; growth regulator; gene; ss.

Brassica rapa.

OS

Key

US2004033606-A1.

19-FEB-2004.

16-AUG-2002; 2002US-00222075.

16-AUG-2002; 2002US-00222075.

(ZOUJ/) ZOU J.

(TAYL/) TAYLOR D C.

(MARI/) MARILLIA E.

Zou J, Taylor DC, Marillia E;

WPI; 2004-24636/23.

P-PSDB; ADL71337.

New purified or recombinant nucleic acid encoding a *Brassica pyruvate*

dehydrogenase kinase protein, for manipulating pyruvate dehydrogenase

activity in plants, e.g. seeds with increased or decreased fatty acid and

oil content.

Claim 2; SEQ ID NO 2; 14bp; English.

The present invention relates to novel *Brassica pyruvate dehydrogenase*

kinase (PDHK) proteins and polynucleotides encoding such proteins.

CC plants resulting in seeds with increased or decreased fatty acid and oil
CC content, plants exhibiting early or delayed flowering time, plants with
CC increased or decreased vegetative growth, plants with root systems better
CC to withstand low soil temperatures or frost, plants with thinner
CC to withstand low soil temperatures or frost, plants with thinner

CC exhibiting higher or lower rates of respiration or plants exhibiting an
 CC enhanced capacity to accumulate biopolymers which rely on acetyl groups
 CC as precursors. They are also useful for generating transgenic plants and
 CC in gene therapy. The present sequence is *Brassica carinata* pyruvate
 CC dehydrogenase kinase (PDHK) cDNA.

XX Sequence 1104 BP; 287 A; 232 C; 294 G; 291 T; 0 U; 0 Other;

Query Match 96.2%; Score 1062.4; DB 12; Length 1104;
 Best Local Similarity 97.6%; Pred. No. 0;
 Matches 1078; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```

QY 1 ATGGCGGTGAAGAGCTGACGAGATGTTTTCGAGAGCTTGATCGAGAGCTTCACAGA 60
DB 1 ATGGCGGTGAAGAGCTGACGAGATGTTTTCGAGAGCTTGATCGAGAGCTTCACAGA 60
QY 61 TGGGGATGCATGAAACGAGAGCGGCTGAGCTCAGGTACATGATGAGATTGCTTCACT 120
DB 61 TGGGGATGCATGAAACGAGAGCGGCTGAGCTCAGGTACATGATGAGATTGCTTCACT 120
QY 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAGAGAGGCTTCGATTCGG 180
DB 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAGAGAGGCTTCGATTCGG 180
QY 181 ATCCGAGAGCGCTGCGATGAACTCGAGACGCTGCTTATGAGCTCTCTGAGAAACCTGCC 240
DB 181 ATCCGAGAGCGCTGCGATGAACTCGAGACGCTGCTTATGAGCTCTCTGAGAAACCTGCC 240
QY 241 GTCTTGAAGTAAAGATGATGATGAGAGTCACTTCAAGGAGCATGAGAGGCTTCCCTGAG 300
DB 241 GTCTTGAAGTAAAGATGATGATGAGAGTCACTTCAAGGAGCATGAGAGGCTTCCCTGAG 300
QY 301 ATCAAGATACCTGCTGATGAGAAAGATTCAACAGATGATCAAGGCTGTTAAAGTAAAG 360
DB 301 ATCAAGATACCTGCTGATGAGAAAGATTCAACAGATGATCAAGGCTGTTAAAGTAAAG 360
QY 361 CACAACAAACGTGGTTCCTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 361 CACAACAAACGTGGTTCCTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 CTCTACGAAAAAGCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 CTCTACGAAAAAGCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 GGGATCCGTATGCTTATGCGGAGCATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 GGGATCCGTATGCTTATGCGGAGCATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 ACAATGGGTATACATACACCAAGATGCTCTTAATGAGAGTGGCAAGAGATGCTATGAA 600
DB 541 ACAATGGGTATACATACACCAAGATGCTCTTAATGAGAGTGGCAAGAGATGCTATGAA 600
QY 601 GATGCAAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 GATGCAAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 GATCCAGATTTACCTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 GATCCAGATTTACCTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 GTCAAGAACTCTGCTCGGTGCTGCAAGAGCGGTTTGTGATGCTGATAGAGGTTGACCA 780
DB 721 GTCAAGAACTCTGCTCGGTGCTGCAAGAGCGGTTTGTGATGCTGATAGAGGTTGACCA 780
QY 781 CCAATCCGATCATTTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 CCAATCCGATCATTTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 GGTGAGGTATACCGAGAGCGGCTCCCTAAATATTAATTAATTAATTAATTAATTAATTA 900
DB 841 GGTGAGGTATACCGAGAGCGGCTCCCTAAATATTAATTAATTAATTAATTAATTAATTA 900
QY 901 AGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960

```

```

DB 901 AGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 CAGATCATATCCATGAGAGATACGGGACTGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 CAGATCATATCCATGAGAGATACGGGACTGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 GACTCGAGAGGCTTGGCATGA 1104
DB 1081 GACTCGAGAGGCTTGGCATGA 1104

```

```

RESULT 4
ADL71334
ID ADL71334 standard; cDNA; 1104 BP.
XX
AC ADL71334;
XX
DT 20-MAY-2004 (first entry)
XX
DE Brassica oleracea pyruvate dehydrogenase kinase (PDHK) cDNA.
XX
KW Pyruvate dehydrogenase kinase; PDHK; fatty acid; oil content; flowering;
XX vegetative growth; respiration; transgenic plant; gene therapy;
XX plant protectant; growth regulator; gene; ss.
XX Brassica oleracea.
XX
OS
XX
FH Key location/Qualifiers
FT 1..1104
FT CDS
FT
FT /product= "PDHK protein"
FT /transl_except= (pos:253..255, aa:Arg)
FT /transl_except= (pos:280..282, aa:Xaa)
FT /note= "Xaa corresponds to Asn or Asp"
FT /transl_except= (pos:316..318, aa:Asp)
FT /transl_except= (pos:592..594, aa:Ala)
FT /transl_except= (pos:622..624, aa:Arg)
FT /transl_except= (pos:652..654, aa:Ile)
FT /transl_except= (pos:886..888, aa:Leu)
FT /transl_except= (pos:952..954, aa:Thr)
FT /transl_except= (pos:979..981, aa:Pro)
XX
PN US2004033606-A1.
XX
PD 19-FEB-2004.
XX
PF 16-AUG-2002; 2002US-00222075.
XX
PR 16-AUG-2002; 2002US-00222075.
XX
PI
XX
PA (ZOU J, ZOU J.
PA (TAYLOR D C.
PA (MARILIA E.
XX
XX
XX Zou J, Taylor DC, Marillia E;
XX
XX WPI, 2004-24636/23.
XX
XX P-PSDB; ADL71338.
XX
XX
XX New purified or recombinant nucleic acid encoding a Brassica pyruvate
XX dehydrogenase kinase protein, for manipulating pyruvate dehydrogenase
XX activity in plants, e.g. seeds with increased or decreased fatty acid and
XX oil content.
XX
XX Claim 2; SEQ ID NO 3; 14bp; English.
XX
XX The present invention relates to novel Brassica pyruvate dehydrogenase
XX kinase (PDHK) proteins and polynucleotides encoding such proteins.
XX

```

CC Sequences of the invention are useful for manipulating PDH activity in
 CC plants resulting in seeds with increased or decreased fatty acid and oil
 CC content, plants exhibiting early or delayed flowering time, plants with
 CC increased or decreased vegetative growth, plants with root systems better
 CC to withstand low soil temperatures or frost, plants with tissues
 CC exhibiting higher or lower rates of respiration or plants exhibiting an
 CC enhanced capacity to accumulate biopolymers which rely on acetyl groups
 CC as precursors. They are also useful for generating transgenic plants and
 CC in gene therapy. The present sequence is *Brassica oleracea* pyruvate
 CC dehydrogenase kinase (PDHK) cDNA.

XX Sequence 1104 BP; 286 A; 229 C; 291 G; 289 T; 0 U; 9 Other;

Query Match 96.2%; Score 1062; DB 12; Length 1104;

Best Local Similarity 97.0%; Pred. No. 0;

Matches 1071; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGGCGGTGAAGAGCTGACGAGATGTTTTCGAAAGCTTGATCGAGACGTTACAGA 60
 DB 1 ATGGCGGTGAAGAGCTGACGAGATGTTTTCGAAAGCTTGATCGAGACGTTACAGA 60
 QY 61 TGGGAGATGATGAAGAGAGGCGGTGAGGCTCAGATGATGATGAGATTCGGTTCAC 120
 DB 61 TGGGAGATGATGAAGAGAGGCGGTGAGGCTCAGATGATGATGAGATTCGGTTCAC 120
 QY 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAAGCTTCCGATTCG 180
 DB 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAAGCTTCCGATTCG 180
 QY 181 ATGCGAGAGCGGTGCGATGCACTTCGAGACGCTGCTTATGAGCTTCTGAGAAACCTGCC 240
 DB 181 ATGCGAGAGCGGTGCGATGCACTTCGAGACGCTGCTTATGAGCTTCTGAGAAACCTGCC 240
 QY 241 GTCCTGAAGGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 DB 241 GTCCTGAAGGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 301 ATCAAGATATCTGCTGATGAGAAAGAGTTTCAACAGATGATCAAGGCTGTTAAAGTAAG 360
 DB 301 ATCAAGATATCTGCTGATGAGAAAGAGTTTCAACAGATGATCAAGGCTGTTAAAGTAAG 360
 QY 361 CACAAACAAGCTGCTCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 DB 361 CACAAACAAGCTGCTCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 421 CTCCTAAGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 DB 421 CTCCTAAGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 481 GGGATCCGATGCTTATCGGCGAGCATGTTGATGATGATGATGATGATGATGATGATGAT 540
 DB 481 GGGATCCGATGCTTATCGGCGAGCATGTTGATGATGATGATGATGATGATGATGATGAT 540
 QY 541 ACAAGTGGTATCATACACCAAGATGCTCTTATGAGAGGTGCAAGAAATGCTATGAA 600
 DB 541 ACTGTGGTATCATACACCAAGATGCTCTTATGAGAGGTGCAAGAAATGCTATGAA 600
 QY 601 GATGCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 DB 601 GATGCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 661 GATGCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 DB 661 GATGCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 QY 721 GTCAAGAACTCTCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 DB 721 GTCAAGAACTCTCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 781 CCAATCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 DB 781 CCAATCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840

QY 841 GGTGAGGTATATACGAGAACGGCTCTCCCTAAATATTTACTTACTTACAGCACTGCA 900
 DB 841 GGTGAGGTATATACGAGAACGGCTCTCCCTAAATATTTACTTACTTACAGCACTGCA 900
 QY 901 AGAACCCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 DB 901 AGAACCCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 QY 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 DB 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 QY 1021 CAGATCATATTCATGAGAGATACGGAGCTGATGATGATGATGATGATGATGATGAT 1080
 DB 1021 CAGATCATATTCATGAGAGATACGGAGCTGATGATGATGATGATGATGATGATGAT 1080
 QY 1081 GACTCGAGAGAGCTTTGCCATGA 1104
 DB 1081 GACTCGAGAGAGCTTTGCCATGA 1104

RESULT 5

AAV49818
 ID AAV49818 standard; cDNA; 1457 BP.

AAV49818;

02-NOV-1998 (first entry)

A. thaliana PDHK cDNA from clone VAS.

Pyruvate dehydrogenase kinase; PDHK; mitochondrion; transgenic plant;

respiration rate; seed oil content; flowering time; cold resistance;

biomass; biopolymer accumulation; fatty acid synthesis; regulation; ss.

Arabidopsis thaliana.

Key Location/Qualifiers

FT CDS 104..1204 /tag= a /product= "PDHK"

FT /note= "pyruvate dehydrogenase kinase"

MO9835044-A1.

13-AUG-1998.

09-FEB-1998; 98MO-CA000096.

10-FEB-1997; 97US-0038815P.

(CANADA) NAT RES COUNCIL CANADA.

Taylor DC, Zou J;

WPI; 1998-447241/38.

P-PSDB; AAME4724.

plant mitochondrial pyruvate dehydrogenase kinase gene - transgenic

plants and seeds; for enhancing biomass, flowering, seed oil content of

soya, corn, cotton.

Claim 1; Fig 2; 86pp; English.

This sequence encodes a mitochondrial pyruvate dehydrogenase kinase,

PDHK, isolated from *Arabidopsis thaliana*. This protein can be used to

produce transgenic plants which have altered respiration rates, seed oil

content and flowering time. Such plants may also have enhanced resistance

to cold, biomass, and capacity to accumulate biopolymers compared to

genetically unmodified plants or seeds. Fatty acid synthesis, vegetative

growth and generation time can also be regulated

Sequence 1457 BP; 395 A; 278 C; 345 G; 439 T; 0 U; 0 Other;

Query Match 79.1%; Score 873.6; DB 2; Length 1457;
 Best Local Similarity 87.6%; Pred. No. 4.7e-262;
 Matches 967; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

1 ATGGCGGTGAAGAGAGCTGAGAGTGTTCGAGAGCTTGATGAGAGAGCTTACAGA 60
 104 ATGGCAGTGAAGAGAGAGCTTCCGAAATGTTCCGAGAGAGTGTGAGAGAGTTCACAA 163
 61 TGGGAGTGTATGAGAGAGAGAGCTGAGAGCTGAGAGTATGAGAGAGTGTTCGACT 120
 164 TGGGAGTGTATGAGAGAGAGAGCTGAGAGCTTATGAGAGAGTGTGAGAGAGTTCGACT 223
 121 CCAGTGTAGAGAGAGAGCTTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 180
 224 CCTATGAGAGAGAGAGCTTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 283
 181 ATGGCAGAGAGAGAGAGAGCTGAGAGAGCTTGTATGAGAGAGTGTTCGACT 240
 284 GTCCGAGAGAGAGAGAGAGCTTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 343
 241 GTCTGAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 300
 344 GTTGTAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 403
 301 ATCAAGAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 360
 404 ATTAAGAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 463
 361 CACAAGAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 420
 464 CATACAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 522
 421 CTACAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 480
 523 --TTCTGAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 580
 481 GGGATCCGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 540
 581 GGGATCCGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 640
 541 ACAGTGGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 600
 641 ACAGTGGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 700
 601 GATGAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 660
 701 GATGAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 760
 661 GATGAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 720
 761 GATGAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 820
 721 GTCAAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 780
 821 GTCAAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 880
 781 CCAATCCGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 840
 881 CCAATCCGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 940
 841 GGTGAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 900
 941 GGTGAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 1000
 901 AGAAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 960
 1001 AGAAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 1060
 961 GGTATGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 1020
 1061 GGTATGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 1120

Query 1021 CAGATCATATCATGAGAGAGATACGGAGAGATGCTTACTGCACTTGTCTGCTTGA 1080
 Db 1121 CAGATCATATCATGAGAGAGATATGGAGAGATGCTTACTGCACTTGTCTGCTTGA 1180
 1081 GACTCCGAGAGAGAGCTTGTCCATGA 1104
 1181 GATTCCGAGAGAGAGCTTATCCCTGA 1204

RESULT 6
 AAD14380
 ID AAD14380 standard; cDNA; 1602 BP.
 AC AAD14380;
 DT 01-NOV-2001 (first entry)
 DE Maize pyruvate dehydrogenase kinase (PDK)-2 cDNA.
 KW Maize; pyruvate dehydrogenase kinase 1; PDK-2; transgenic plant; ss.
 OS Zea mays.
 FH Key location/Qualifiers
 FT CDS /tag= a /product= "Maize PDK-2 protein"
 FT sig_peptide /tag= b /product= "Maize PDK-2 protein"
 FT mat_peptide /tag= c /product= "Maize PDK-2 protein"
 FT FT
 XX US6265636-B1.
 XX PD 24-JUL-2001.
 XX PF 15-JUN-1999; 99US-00333423.
 XX PR 19-JUN-1998; 98US-0089998P.
 XX PA (PION-) PIONEER HI-BRED INT INC.
 XX PA (UMOR) UNIV MISSOURI.
 XX PI Randall DD, Thelen J, Miernyk JA, Muszynski MG, Sewalt VH;
 DR WPI; 2001-463947/50.
 DR P-PSDB; AAE07838.
 XX PT Nucleic acids encoding monocot pyruvate dehydrogenase kinase, useful for
 PT producing plants, e.g. corn cotton and rice with altered characteristics
 PT e.g. oil content.
 XX PS Claim 1b; Col 61-64; 36pp; English.
 XX CC The present sequence is a cDNA coding for maize pyruvate dehydrogenase
 CC kinase (PDK)-2. The PDK cDNAs are used via standard recombinant DNA
 CC methodologies for the production of transgenic plants (e.g. corn,
 CC soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley
 CC and/or millet) which have altered characteristics and metabolic
 CC processes, e.g. altered carbon flow into the Krebs cycle and/or
 CC increased oil content
 XX SQ Sequence 1602 BP; 388 A; 395 C; 428 G; 391 T; 0 U; 0 Other;

Query Match 47.5%; Score 524.6; DB 5; Length 1602;
 Best Local Similarity 68.8%; Pred. No. 7.7e-153;
 Matches 760; Conservative 0; Mismatches 324; Indels 21; Gaps 2;

12 GAAGCTAGAGAGAGTGTTCGAGAGAGCTTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 71
 77 GATGAGAGAGAGAGAGAGTGTATGAGAGAGTGTTCGAGAGAGTGTTCGACT 136


```

Qy 72 GAAGCAGACGGGCGTGAAGCTCAGGTACATGATGAGATTGCGTTCCATCCCACTGAGAG 131
Db 137 GAAGCAGACGGGCGGTTGACCTCGCGGTACATGATGAGATTGCGTTCCCGCCCAACCGAGCG 196
Qy 132 AAACCTTCTGATCTGCGCGCAGTTTCTTCAAGAGAGCTTCGATTGCGATGCGGAGCG 191
Db 197 CAACCTCTCTCTCCCGCAGTTCTTGCAACAGAGAGCTCCCACTGCGCTTCCGCGCGCG 256
Qy 192 TGGGATGGAATCGGAGAGCGCTCTTATGCGCTCTTGAGAAACCTCCCGCTTGAAGT 251
Db 257 CGGCTGAGCTGAGCTGCTGCGCTTCCGCTCTTCCAAAGAGCGCGCATCTTCAAGT 316
Qy 252 AAGAGATTGATGATGAGATCATTCAGGAGCATGAGAGCGTTTCTGATCAAGATAC 311
Db 317 GCGGAGCTGGTACTTGAGCTCATTCGCGGACATCAGATTCTTCCCTGAAGTGAAGACCG 376
Qy 312 TGCTGATGAGAAAGATTACACAGATGATCAAGGCTTTAAGTAAAGGCAACAAGT 371
Db 377 GAAAGAGAGCTCTCTTTCACGAGATGATCAATATGATCAAGGCTGCGGCAATCAATGT 436
Qy 372 GATTCCATGATGCTGCGGTGTAACCAAGCTGAAGAAAGATGAACCTTACGAAA 431
Db 437 GGTTCACACCACTGCTTGGAGTGCAGAGCTGAAGAGAGAGCTGCGCGCTCAAGAA 496
Qy 432 GCTT-----GATGATTCATCAGTTTCTTGATGCTTCTTACTTGTCTGAT 479
Db 497 GGTTCATTCAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 556
Qy 480 AAGGATTCGATGCTTATTCGCGGAGCATGTTGATGCTGATCAATCAACCAACCATTTCA 539
Db 557 TGGCATCCGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 616
Qy 540 CACAGTGGGTTTACATACACCAAGATGCTCTTATGAGAGTGGCAAGATGATGATGAT 599
Db 617 TGTATTAGGCTCTACATACCAAGATGATGATGATGATGATGATGATGATGATGAT 676
Qy 600 AGATGCAAGGTCGATTTGTTTTCAGAGATGATGATGATGATGATGATGATGATGATGAT 659
Db 677 AGATGCGCGTTCGTCTGTTTAAAGGATGATGATGATGATGATGATGATGATGATGAT 736
Qy 660 CGATCCAGATTTCATTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 719
Db 737 CGATCCAAATTTTACGTTTCCATGATGATGATGATGATGATGATGATGATGATGATGAT 796
Qy 720 AGTCAAGATCTCTCGGTGCTGTCGAAGAGCGGTTTGTGATCTGATGATGATGATGATGAT 779
Db 797 GGTGAAGATCTCTCGCGCGGTGTCGAAGAGCGGTTTGTGATCTGATGATGATGATGATGAT 856
Qy 780 ACCAATCCGATCATTTGTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 839
Db 857 CCGGTTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 916
Qy 840 AGTGAAGTATACCGAAGAGCGGTTCTCTAAATTAATCACTTACCTTACAGACTGC 899
Db 917 AGGTGGGCGGATACCAAGAGCGGCTCCGAGATTTTCAAGATCTGATGATGATGATGATGAT 976
Qy 900 AAGAAACCCACTGAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959
Db 977 AAGAAATCTCTGAGCTGATGAT-----AGCTTAATTAATGAAAGGATGATGATGATGAT 1027
Qy 960 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019
Db 1028 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087
Qy 1020 GCAAGATCATATCCATGAGAGATGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1079
Db 1088 GCAAAATCATCTGATGAGAGATGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1147
Qy 1080 AGACTGCGAGAGCGCTTGCATGA 1104
Db 1148 AGATTCAAGAGAGCGCTTGCCTTGA 1172

```

```

RESULT 7
AADI4379
ID AADI4379 standard; cDNA; 1332 BP.
XX
AC AADI4379;
XX
DT 01-NOV-2001 (first entry)
XX
DE Maize pyruvate dehydrogenase kinase (PDK)-1 cDNA.
XX
KW Maize; pyruvate dehydrogenase kinase 1; PDK-1; transgenic plant; ss.
XX
OS Zea mays.
XX
FH Key
FT CDS Location/Qualifiers
FT FT /*tag= a
FT FT /product= "Maize PDK-1 protein"
FT FT sig_peptide /*tag= b
FT FT mat_peptide /*tag= b
FT FT /*tag= c
FT FT /product= "Mature maize PDK-1 protein"
XX
PN US6265636-B1.
PD 24-JUL-2001.
XX
PE 15-JUN-1999; 99US-00333423.
XX
PR 19-JUN-1998; 98US-0089998P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
PA (UMOR ) UNIV MISSOURI.
PI Randall DD, Thelen JF, Mierznyk JA, Muszynski MG, Sewalt VH;
XX
XX WPI; 2001-463947/50.
DR P-PSDB; AAB07837.
XX
XX Nucleic acids encoding monocot pyruvate dehydrogenase kinase, useful for
PT producing plants, e.g. corn cotton and rice with altered characteristics
PT e.g. oil content.
XX
PS Claim 1b; Col 57-60; 36pp; English.
XX
CC The present sequence is a cDNA coding for maize pyruvate dehydrogenase
CC kinase (PDK)-1. The PDK cDNAs are used via standard recombinant DNA
CC methodologies for the production of transgenic plants (e.g. corn,
CC soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley
CC and/or millet) which have altered characteristics and metabolic
CC processes, e.g. altered carbon flow into the Krebs cycle and/or
CC increased oil content
XX
SQ Sequence 1332 BP; 345 A; 311 C; 345 G; 331 T; 0 U; 0 Other;
Query Match 40.6%; Score 448.2; DB 5; Length 1332;
Best Local Similarity 66.1%; Pred. No. 5; se-129;
Matches 732; Conservative 0; Mismatches 303; Indels 72; Gaps 3;
Qy 10 AAGAGGCTGACGAGATGTTTTCGAAGAGCTTATGATGAGAGCTTCAAGATGAGGATGC 69
Db 52 AGGATGCGCGTGGAGCGCGGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 111
Qy 70 ATGAAGCAGACGGGCGTGAAGCTTCAGATGATGATGATGATGATGATGATGATGATGATGAT 129
Db 112 ATGCGCAGACCGGCGGTGACCTGCGCTACATGATGATGATGATGATGATGATGATGATGAT 171
Qy 130 AGAACTTCTGATCTCGGCGCAGTTTCTTCAAGAGAGCTTCCGATGCGATGCGAGG 189
Db 172 CGGACCTCTCTTCTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 231
Qy 190 CGTGCATGCACTCGAGAGCGCTGCTTATGCGCTCTCTGAGAAACCTGCGCTTGAAG 249

```


Qy	428	AAAAGCTGATAGATATCAAGTTCTTGATGCGTTCACTGTCCGATAGGGATCC	487
Db	62	AGGATCTTGATAGATTCATCAAGTTCTTCGATGGTTTACTTGTCAAGATTGGAAATCC	121
Qy	488	GTAATCTTATCGGGCAGCATGTTGAGTTCATATCAAAACCAACAATTCCACACAGTGG	547
Db	122	GCATCTTATTTGGGGCAGCAGTCGACCTCATTAACCAATCCTGCTCCACATGTTGTGG	181
Qy	548	GTTACATACACCAAGATGTCTCCTATGAGAGGTGCAGGAATGCTAGTGAATGSCAA	607
Db	182	GCTATATATCATCAAAAATGCTCTCAGTGAAGGTGGCAAAAAATGCTAGGAGGCTGCAC	241
Qy	608	GGTGCAATTTGTTTCAGAGAGTATGTTCTGCTTCGAGATTAACATATATGGCGATCCAA	667
Db	242	GTGCTATTTGTTGGCTGATATGGAAGTGCACCTGATATATATCTAAGGTATCTTA	301
Qy	668	GTTTCACTTTCCGATATGTTCCGACCCCATTTGCATCTTATNGGTATAGATTAGTCAAGA	727
Db	302	GTTTACATTTCTTACGTTCAACAACAACCTTGACACTTATGTAATTCGAGCTGTCAAAA	361
Qy	728	ACTCTCTCCGTGCTGCCAAGAGCGGTTTGTGACTCTGATAGAGTTGCACCAACAATCC	787
Db	362	ATTCCTGTGCTGTGCCAAGAGCGCTTTATGTAATCAATTAAGTTGCACCACTGTTTC	421
Qy	788	GTAATCATTTGTTCTGTATGTAATTCGAAGATGTTTCAATTAAGGTTCTCAGATGAAGGTGGAG	847
Db	422	GAATATATAGTTGCCAGAGGGAATCGAGATGTAACTTAAAGATCTCAGATGAAGGGGGTG	481
Qy	848	GTAATCCGAGAGCGGTCTCCCTAAATAATTCCTAACCTCTACAGCACTGCAGAAAGAC	907
Db	482	GCATTAAGAGAGTGGTCTTCCCAAAATTTTCACTTACCTTTATAGACCTGCCAAAAAC	541
Qy	908	CACCTGAAGAGATGTGACCTTGGGAACCGCTGAT	942
Db	542	CCTTGATGAGCATACAGATCTTGAACAGATGAT	576

XX	RESULT 9
XX	ADM80480
ID	ADM80480 standard; DNA, 584 BP.
AC	
XX	ADM80480;
DT	01-JUL-2004 (first entry)
XX	
DE	Maize cDNA #1 used in translation initiation codon detection method.
XX	
KW	Translation initiation codon determination; ATG codon;
KW	Quadratic Discriminant Analysis; QDA; initiator codon;
XX	pseudoinitiator codon; ATG triplet; bioinformatic; maize; ss.
OS	Zea mays.
XX	
PN	US2004067514-A1.
XX	
PD	08-APR-2004.
XX	
PF	16-JUL-2003; 2003US-00620796.
XX	
PR	17-JUL-2002; 2002US-0396908P.
XX	
PA	(TABA/) TABASKA J E.
XX	
PI	Tabaska JE;
DR	WI; 2004-355292/33.
XX	
PT	Finding translation initiation codons, useful as bioinformatics tool for
PT	analyzing files of nucleic acid sequence data, comprises using Quadratic
PT	Discriminant Analysis to determine the translation initiation codon.
XX	
PS	Example 1; SEQ ID NO 1; 19pp; English.
XX	

CC The present invention relates to a method of finding translation
 CC initiation codons in a nucleotide sequence using Quadratic Discriminant
 CC Analysis (QDA) to determine the translation initiation (ATG) codon in a
 CC nucleotide sequence. The method comprises (a) analysing a first data set
 CC to measure a combination of features of initiator codons and
 CC pseudoinitiator codons and to produce a set of numerical values for the
 CC combination of features, (b) evaluating scoring functions by reading a
 CC sequence in the vicinity of an ATG triplet and using the scoring
 CC functions and the scoring function's parameters to return a numerical
 CC score that quantifies how much the ATG triplet resembles an initiator
 CC codon, (c) generating a quadratic discriminant function through selection
 CC of a combination of feature variables that optimally classifies ATG
 CC triplets in a nucleotide sequence as initiator codons or as
 CC pseudoinitiator codons based on the output of the scoring functions and
 CC through the use of Quadratic Discriminant Analysis, and (d) using the
 CC quadratic discriminant function to analyse a second data set of
 CC nucleotide sequences by evaluating at least one scoring function for each
 CC ATG triplet in the sequences and to calculate the probability of an
 CC initiator codon at a position using the output of the analysis. The
 CC method is useful for finding translation initiation codons in a
 CC nucleotide sequence. The method is useful as a bioinformatics tool for
 CC analysing files of nucleic acid sequence data to find translation
 CC initiation codons. The present sequence represents a *maize* cDNA sequence
 CC used in the exemplification of the present invention.

XX Sequence 584 BP; 117 A; 167 C; 191 G; 109 T; 0 U; 0 Other;

Query Match	22.18;	Score 244.2;	DB 12;	Length 584;
-------------	--------	--------------	--------	-------------

Matches 361; Conservative 0; Mismatches 158; Indels 12; Gaps 1.

QY	12	GAAGGCTAGACGAGATGTTTTGAAAGACCTTATCAGAGACCTTCACAATGGGGATGCAAT	71
Db	50	GATGGCCTCGAGGCGGCTGCGGGGGGGGTGGCGAGGAGGATGGGCGCGTGGGGACGACAT	109
QY	72	GAAGCAGACGGGCGTGAACCTCAGGTACATGATGGAATTCGGTTCCATCCCATGAGAG	131
Db	110	GAAGCACAACGGGGGTGACCCTCGCGGTACATATGAGATTTGGACTCCCGCCCAACCGAGG	169
QY	132	AAACCTTCTGATCTCGGGGAGTTTCTTCAAGAGAGCTTCGGATTCCGATTCGAGATCGAGGCG	191
Db	170	AAACCTGCTCCTCTTCGGGCGATTCTGACAAAGAGCTCCCATCCGCTTGGCACGCG	229
QY	192	TGCGATGAACTCGAGACGCTGCTCTTATGCGCTTCTTGAGAAACCTGCGCTTGAAGT	251
Db	230	CGCGCTCGAGCTCGACTCGCTGCGCCTTCGCGCTCTCAACAAGCCGCGCATCTCAAGGT	289
QY	252	AAGGATTTGGATTTGGATTTGAGTCAATCAGGSAATGAGACGTTTCTTGATTAAGATAC	311
Db	290	GCGGGACTGTGACTTTGACTATTTCCGGGACATCAGATCACTTCCCTGAAGTGAAGAGCCG	349
QY	312	TGCTGATGAGAAAGATTACACAGATGATCAAGGCTTTAAAGTAAAGGCAACAAACGT	371
Db	350	GAAGAGAGAGTGGCTTTCAACGAGATGATCAATATGGTCAAGGTGGGCAATAAACAATGT	409
QY	372	GATTCCATGATGGCTCTGGGTGTGAACCACTGAAAGAAAGAAATGAAACTTACGAAA	431
Db	410	GATTCCAACCATGGCCTTGGGAGTGCACAGCACTGAAGAAAGAGCTGGGCGGTCAAGAA	469
QY	432	GCTT-----GATGAGATTCACTGATTTCTTATGCTTCACTTGTCTCTGAT	479
Db	470	GATTCCATTCGAATTCGATGAGATCGACGATTTCTTTGACCGGTTCTTCAATGTCAAGAA	529
QY	480	AGGATCCGATAGCTTATCGGGCAGCATGTTGATGCTATTAATCCAAACC	530
Db	530	TGGCATTCGCAATGCTGATAGGGCAGCAATGTGGCTTTGATATACCTTAACC	580

RESULT	10
AAC36807	
ID	AAC36807 standard; DNA; 464 BP
XX	
AC	AAC36807;

XX	17-OCT-2000 (first entry)		PR	18-JUN-1999;	99US-0139750P
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 15144.		PR	18-JUN-1999;	99US-0139763P
DE	Hybridisation assay; genetic mapping; gene expression control;		PR	21-JUN-1999;	99US-0139817P
XX	protein identification; signal transduction pathway; metabolic pathway;		PR	22-JUN-1999;	99US-0139899P
KM	Promoter; termination sequence; ss.		PR	23-JUN-1999;	99US-0140353P
XX	Arabidopsis thaliana.		PR	23-JUN-1999;	99US-0140354P
OS	BP1033405-A2.		PR	24-JUN-1999;	99US-0140695P
XX	06-SEP-2000.		PR	28-JUN-1999;	99US-0140823P
PD	25-FEB-2000; 2000EP-00301439.		PR	29-JUN-1999;	99US-0140991P
XX	25-FEB-1999;	99US-0121825P.	PR	30-JUN-1999;	99US-0141287P
PR	05-MAR-1999;	99US-0123180P.	PR	01-JUL-1999;	99US-0141842P
PR	09-MAR-1999;	99US-0123548P.	PR	02-JUL-1999;	99US-0142154P
PR	23-MAR-1999;	99US-0125588P.	PR	06-JUL-1999;	99US-0142390P
PR	29-MAR-1999;	99US-0126284P.	PR	08-JUL-1999;	99US-0142803P
PR	29-MAR-1999;	99US-0126785P.	PR	09-JUL-1999;	99US-0142920P
PR	01-APR-1999;	99US-0127462P.	PR	12-JUL-1999;	99US-0143297P
PR	06-APR-1999;	99US-0128234P.	PR	13-JUL-1999;	99US-0143342P
PR	08-APR-1999;	99US-0128714P.	PR	14-JUL-1999;	99US-0143624P
PR	16-APR-1999;	99US-0129845P.	PR	15-JUL-1999;	99US-0144005P
PR	19-APR-1999;	99US-0130077P.	PR	16-JUL-1999;	99US-0144085P
PR	21-APR-1999;	99US-0130449P.	PR	16-JUL-1999;	99US-0144086P
PR	23-APR-1999;	99US-0130510P.	PR	19-JUL-1999;	99US-0144325P
PR	28-APR-1999;	99US-0130891P.	PR	19-JUL-1999;	99US-0144331P
PR	30-APR-1999;	99US-0131449P.	PR	20-JUL-1999;	99US-0144332P
PR	30-APR-1999;	99US-0132048P.	PR	20-JUL-1999;	99US-0144632P
PR	04-MAY-1999;	99US-0132407P.	PR	21-JUL-1999;	99US-0144814P
PR	05-MAY-1999;	99US-0132484P.	PR	21-JUL-1999;	99US-0145086P
PR	06-MAY-1999;	99US-0132485P.	PR	21-JUL-1999;	99US-0145088P
PR	06-MAY-1999;	99US-0132486P.	PR	22-JUL-1999;	99US-0145087P
PR	07-MAY-1999;	99US-0132487P.	PR	22-JUL-1999;	99US-0145193P
PR	11-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145194P
PR	14-MAY-1999;	99US-0134218P.	PR	23-JUL-1999;	99US-0145216P
PR	14-MAY-1999;	99US-0134219P.	PR	26-JUL-1999;	99US-0145224P
PR	14-MAY-1999;	99US-0134221P.	PR	26-JUL-1999;	99US-0145276P
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145913P
PR	18-MAY-1999;	99US-0134376P.	PR	27-JUL-1999;	99US-0145918P
PR	19-MAY-1999;	99US-0134941P.	PR	28-JUL-1999;	99US-0145951P
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146386P
PR	21-MAY-1999;	99US-0135253P.	PR	02-AUG-1999;	99US-0146388P
PR	24-MAY-1999;	99US-0135629P.	PR	03-AUG-1999;	99US-0147038P
PR	25-MAY-1999;	99US-0136021P.	PR	04-AUG-1999;	99US-0147204P
PR	27-MAY-1999;	99US-0136392P.	PR	05-AUG-1999;	99US-0147192P
PR	28-MAY-1999;	99US-0136782P.	PR	06-AUG-1999;	99US-0147260P
PR	01-JUN-1999;	99US-0137222P.	PR	06-AUG-1999;	99US-0147303P
PR	03-JUN-1999;	99US-0137528P.	PR	09-AUG-1999;	99US-0147416P
PR	04-JUN-1999;	99US-0137500P.	PR	09-AUG-1999;	99US-0147935P
PR	07-JUN-1999;	99US-0137722P.	PR	10-AUG-1999;	99US-0148171P
PR	08-JUN-1999;	99US-0138094P.	PR	11-AUG-1999;	99US-0148319P
PR	10-JUN-1999;	99US-0138540P.	PR	13-AUG-1999;	99US-0148341P
PR	10-JUN-1999;	99US-0138847P.	PR	13-AUG-1999;	99US-0148563P
PR	14-JUN-1999;	99US-0139119P.	PR	16-AUG-1999;	99US-0148684P
PR	16-JUN-1999;	99US-0139452P.	PR	16-AUG-1999;	99US-0149368P
PR					

```

PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 10-SEP-1999; 99US-0152363P.
PR 13-SEP-1999; 99US-0153070P.
PR 15-SEP-1999; 99US-0153758P.
PR 16-SEP-1999; 99US-0154018P.
PR 20-SEP-1999; 99US-0154039P.
PR 22-SEP-1999; 99US-0154779P.
PR 23-SEP-1999; 99US-0155139P.
PR 24-SEP-1999; 99US-0155486P.
PR 28-SEP-1999; 99US-0155659P.
PR 29-SEP-1999; 99US-0156458P.
PR 04-OCT-1999; 99US-0156596P.
PR 05-OCT-1999; 99US-0157117P.
PR 06-OCT-1999; 99US-0157753P.
PR 07-OCT-1999; 99US-0157865P.
PR 08-OCT-1999; 99US-0158029P.
PR 12-OCT-1999; 99US-0158232P.
PR 13-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161992P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      21.4%; Score 236.2; DB 3; Length 464;
Best Local Similarity 83.5%; Pred. No. 6,7e-63;
Matches 268; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

```

```

DB 384 GTTTGAGGTGCGTATGATTTGGAATCTTTCAGGACATGAGACATTTCCGAG 443
OY 301 ATCAGAGATTCCTGATGAG 321
DB 444 ATTAGGATTCGGGTGACGAG 464

```

RESULT 11

```

ID AAA79663 standard; cDNA; 371 BP.
XX AAA79663;
AC AAA79663;
XX 27-NOV-2000 (first entry)
DT 27-NOV-2000 (first entry)
DE Eucalyptus grandis cell signalling involved polynucleotide SEQ ID NO:464.
XX Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
KM plant cell signalling; modulation; transgenic plant; pathogen; growth;
KM environmental change; development; cell proliferation; differentiation;
XX elongation; survival; disease resistance; nutrient metabolism; ss.
XX Eucalyptus grandis.
OS Eucalyptus grandis.
XX MO200042171-A1.
PN 20-JUL-2000.
PD 20-JUL-2000.
XX 11-JAN-2000; 2000MO-US000724.
PF 11-JAN-2000; 2000MO-US000724.
XX 12-JAN-1999; 99US-00228986.
PR 01-NOV-1999; 99US-0162866P.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX Strabala TJ, Nieuwenhuizen NJ;
XX WPI; 2000-476052/41.
XX Isolated polynucleotide encoding a polypeptide involved in cell signaling
PT used for generating transgenic plants with modified responses to external
PT signals.
XX Claim 1; Page 208; 527pp; English.
XX AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
CC pine (Pinus radiata also known as Monterey pine). The protein sequences
CC are involved in cell signalling. The polynucleotide and protein sequences
CC can be used to modify the response of plant cells to external signals
CC e.g. environmental changes or pathogens during the growth and development
CC of a plant. They can be used to modify cell proliferation,
CC differentiation, elongation and survival, resistance to disease and
CC nutrient metabolism. Examples of modifications which can be produced are
CC altered fruit ripening and senescence of leaves and flowers e.g. to delay
CC senescence and prolong the life of cut flowers or enhance senescence of
CC reproductive organs to engineer sterile plants. Other modifications can
CC be used to delay senescence in selected cell types or organs providing
CC fruit and vegetables which have a longer shelf life between harvest and
CC consumption, or to decrease branching frequency in forest tree species
CC giving long stretches of valuable knot-free clear wood which can be used
CC in solid timber furniture and veneers
XX
SQ Sequence 371 BP; 97 A; 74 C; 93 G; 107 T; 0 U; 0 Other;
Query Match      18.9%; Score 208.2; DB 3; Length 371;
Best Local Similarity 76.8%; Pred. No. 3.4e-54;
Matches 268; Conservative 0; Mismatches 78; Indels 3; Gaps 1;
OY 697 TTGCATCTTATGCTATGATGATTAGTCAAGAACTCTTCGCTGCTGTCCAGAGCGGTTT 756
DB 2 TTGCATCTTATGCTATGATGATTAGTCAAGAACTCTTCGCTGCTGTCCAGAGCGGTTT 61

```

CC having improved properties comprising transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX

SQ Sequence 1523 BP; 385 A; 461 C; 368 G; 309 T; 0 U; 0 Other;

```

RESULT 12
ADS49235
ID ADS49235 standard; cDNA; 1523 BP.
vv

```

Bacterial polynucleotide #33978.

KM recombinant DNA construct; transformed plant; improved plant property;
KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KM pathogen tolerance; pest tolerance; plant disease resistance;
KM cell cycle pathway modification; plant disease regulator;
KM homologous recombination; seed oil yield; protein yield; carbohydrate;
KM nitrogen; phosphorus; photosynthesis; lignin; galactomanan;
KM bacterial polynucleotide; gene; ss.

PN US2003233675-A1.

PF 20-FEB-2003; 2003US-00369493.

PR 21-FEB-2002; 2002US-0360039P.

PA (CAOY/) CAO Y.

PA (SLAT/) SLATER S C.

PA (GOLD/) GOLDMAN B S.

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman RS.

DR WPI; 2004-061375/06.

PT New recombinant DNA construct comprising a promoter positioned to provide

microbial source, useful for producing plants with improved properties

PS Claim 1; SEQ ID NO 27665: 122bp: English.

CC The invention relates to a recombinant DNA construct comprising a

provide for expression of a polynucleotide encoding a polypeptide from a promoter functional in a plant cell, where the promoter is positioned to

comprising the recombinant DNA construct and a method of producing a

such as maize or soybean. The method of producing a transformed plant

326 AGTTTCAACAAGATGATTAAAGGAATCTATAACATAACCCATAAACAACATC
OY

[illegible][illegible]

728

416

780

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

1. The first group of respondents (n = 10) was asked to rate the importance of each of the 10 factors on a scale of 1 to 5, where 1 was 'not important at all' and 5 was 'very important'. The results are shown in Table 1.

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

A vertical ruler with markings from 0 to 10 cm. The markings are in centimeters, with millimeter increments indicated between the centimeter numbers. The ruler is oriented vertically, with the 0 mark at the top and the 10 cm mark at the bottom.

100T 202Y 203Y 204Y 205Y 206Y 207Y 208Y 209Y 210Y 211Y 212Y 213Y 214Y 215Y 216Y 217Y 218Y 219Y 220Y 221Y 222Y 223Y 224Y 225Y 226Y 227Y 228Y 229Y 230Y 231Y 232Y 233Y 234Y 235Y 236Y 237Y 238Y 239Y 240Y 241Y 242Y 243Y 244Y 245Y 246Y 247Y 248Y 249Y 250Y 251Y 252Y 253Y 254Y 255Y 256Y 257Y 258Y 259Y 260Y 261Y 262Y 263Y 264Y 265Y 266Y 267Y 268Y 269Y 270Y 271Y 272Y 273Y 274Y 275Y 276Y 277Y 278Y 279Y 280Y 281Y 282Y 283Y 284Y 285Y 286Y 287Y 288Y 289Y 290Y 291Y 292Y 293Y 294Y 295Y 296Y 297Y 298Y 299Y 300Y 301Y 302Y 303Y 304Y 305Y 306Y 307Y 308Y 309Y 310Y 311Y 312Y 313Y 314Y 315Y 316Y 317Y 318Y 319Y 320Y 321Y 322Y 323Y 324Y 325Y 326Y 327Y 328Y 329Y 330Y 331Y 332Y 333Y 334Y 335Y 336Y 337Y 338Y 339Y 340Y 341Y 342Y 343Y 344Y 345Y 346Y 347Y 348Y 349Y 350Y 351Y 352Y 353Y 354Y 355Y 356Y 357Y 358Y 359Y 360Y 361Y 362Y 363Y 364Y 365Y 366Y 367Y 368Y 369Y 370Y 371Y 372Y 373Y 374Y 375Y 376Y 377Y 378Y 379Y 380Y 381Y 382Y 383Y 384Y 385Y 386Y 387Y 388Y 389Y 390Y 391Y 392Y 393Y 394Y 395Y 396Y 397Y 398Y 399Y 400Y 401Y 402Y 403Y 404Y 405Y 406Y 407Y 408Y 409Y 410Y 411Y 412Y 413Y 414Y 415Y 416Y 417Y 418Y 419Y 420Y 421Y 422Y 423Y 424Y 425Y 426Y 427Y 428Y 429Y 430Y 431Y 432Y 433Y 434Y 435Y 436Y 437Y 438Y 439Y 440Y 441Y 442Y 443Y 444Y 445Y 446Y 447Y 448Y 449Y 450Y 451Y 452Y 453Y 454Y 455Y 456Y 457Y 458Y 459Y 460Y 461Y 462Y 463Y 464Y 465Y 466Y 467Y 468Y 469Y 470Y 471Y 472Y 473Y 474Y 475Y 476Y 477Y 478Y 479Y 480Y 481Y 482Y 483Y 484Y 485Y 486Y 487Y 488Y 489Y 490Y 491Y 492Y 493Y 494Y 495Y 496Y 497Y 498Y 499Y 500Y 501Y 502Y 503Y 504Y 505Y 506Y 507Y 508Y 509Y 510Y 511Y 512Y 513Y 514Y 515Y 516Y 517Y 518Y 519Y 520Y 521Y 522Y 523Y 524Y 525Y 526Y 527Y 528Y 529Y 530Y 531Y 532Y 533Y 534Y 535Y 536Y 537Y 538Y 539Y 540Y 541Y 542Y 543Y 544Y 545Y 546Y 547Y 548Y 549Y 550Y 551Y 552Y 553Y 554Y 555Y 556Y 557Y 558Y 559Y 560Y 561Y 562Y 563Y 564Y 565Y 566Y 567Y 568Y 569Y 570Y 571Y 572Y 573Y 574Y 575Y 576Y 577Y 578Y 579Y 580Y 581Y 582Y 583Y 584Y 585Y 586Y 587Y 588Y 589Y 590Y 591Y 592Y 593Y 594Y 595Y 596Y 597Y 598Y 599Y 600Y 601Y 602Y 603Y 604Y 605Y 606Y 607Y 608Y 609Y 610Y 611Y 612Y 613Y 614Y 615Y 616Y 617Y 618Y 619Y 620Y 621Y 622Y 623Y 624Y 625Y 626Y 627Y 628Y 629Y 630Y 631Y 632Y 633Y 634Y 635Y 636Y 637Y 638Y 639Y 640Y 641Y 642Y 643Y 644Y 645Y 646Y 647Y 648Y 649Y 650Y 651Y 652Y 653Y 654Y 655Y 656Y 657Y 658Y 659Y 660Y 661Y 662Y 663Y 664Y 665Y 666Y 667Y 668Y 669Y 670Y 671Y 672Y 673Y 674Y 675Y 676Y 677Y 678Y 679Y 680Y 681Y 682Y 683Y 684Y 685Y 686Y 687Y 688Y 689Y 690Y 691Y 692Y 693Y 694Y 695Y 696Y 697Y 698Y 699Y 700Y 701Y 702Y 703Y 704Y 705Y 706Y 707Y 708Y 709Y 710Y 711Y 712Y 713Y 714Y 715Y 716Y 717Y 718Y 719Y 720Y 721Y 722Y 723Y 724Y 725Y 726Y 727Y 728Y 729Y 730Y 731Y 732Y 733Y 734Y 735Y 736Y 737Y 738Y 739Y 740Y 741Y 742Y 743Y 744Y 745Y 746Y 747Y 748Y 749Y 750Y 751Y 752Y 753Y 754Y 755Y 756Y 757Y 758Y 759Y 760Y 761Y 762Y 763Y 764Y 765Y 766Y 767Y 768Y 769Y 770Y 771Y 772Y 773Y 774Y 775Y 776Y 777Y 778Y 779Y 780Y 781Y 782Y 783Y 784Y 785Y 786Y 787Y 788Y 789Y 790Y 791Y 792Y 793Y 794Y 795Y 796Y 797Y 798Y 799Y 800Y 801Y 802Y 803Y 804Y 805Y 806Y 807Y 808Y 809Y 810Y 811Y 812Y 813Y 814Y 815Y 816Y 817Y 818Y 819Y 820Y 821Y 822Y 823Y 824Y 825Y 826Y 827Y 828Y 829Y 830Y 831Y 832Y 833Y 834Y 835Y 836Y 837Y 838Y 839Y 840Y 841Y 842Y 843Y 844Y 845Y 846Y 847Y 848Y 849Y 850Y 851Y 852Y 853Y 854Y 855Y 856Y 857Y 858Y 859Y 860Y 861Y 862Y 863Y 864Y 865Y 866Y 867Y 868Y 869Y 870Y 871Y 872Y 873Y 874Y 875Y 876Y 877Y 878Y 879Y 880Y 881Y 882Y 883Y 884Y 885Y 886Y 887Y 888Y 889Y 890Y 891Y 892Y 893Y 894Y 895Y 896Y 897Y 898Y 899Y 900Y 901Y 902Y 903Y 904Y 905Y 906Y 907Y 908Y 909Y 910Y 911Y 912Y 913Y 914Y 915Y 916Y 917Y 918Y 919Y 920Y 921Y 922Y 923Y 924Y 925Y 926Y 927Y 928Y 929Y 930Y 931Y 932Y 933Y 934Y 935Y 936Y 937Y 938Y 939Y 940Y 941Y 942Y 943Y 944Y 945Y 946Y 947Y 948Y 949Y 950Y 951Y 952Y 953Y 954Y 955Y 956Y 957Y 958Y 959Y 960Y 961Y 962Y 963Y 964Y 965Y 966Y 967Y 968Y 969Y 970Y 971Y 972Y 973Y 974Y 975Y 976Y 977Y 978Y 979Y 980Y 981Y 982Y 983Y 984Y 985Y 986Y 987Y 988Y 989Y 990Y 991Y 992Y 993Y 994Y 995Y 996Y 997Y 998Y 999Y 1000Y 1001Y 1002Y 1003Y 1004Y 1005Y 1006Y 1007Y 1008Y 1009Y 1010Y 1011Y 1012Y 1013Y 1014Y 1015Y 101

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

...GAGACATCACCATCCTGGATGAAGGAGCGGATTC 1207

*7
001 CCGGAGCCGCCTCCTTATTTATTCACCTAACGCACTGGCAAGAACCACATTG 913

1208 CCCGAGCGCAI CCCCGTIGTIGGACCTACAGTACACCA CCGTCGATCGCACACCCA 1267

514 AAGAGAGAGGACCTGGGACCGCTGAGTGTCCCTGACCTATGGCTGGTTATGGTATG 973

1208 ACCGGATCCCGACTTCGACAGAGCGACCTTAAGGCCCCCATGGCCGGTTCCGCTACG 1327

9/4 GTCCTGCTTGTATGCTGGCTATTTGGTGAGATTGCAGATCATATCCA 1033

1328 GACCTTCCTATCTCGCGCTGTATCGGAGATACTTGGTGGCGACCTTAAGTTGATCAGCA 1387

OS Aspergillus fumigatus.
 XX WO200286090-A2.
 XX 31-OCT-2002.
 PD
 XX 23-APR-2002; 2002WO-US013142.
 PF
 XX 23-APR-2001; 2001US-0285697P.
 PR 27-APR-2001; 2001US-0287066P.
 PR 05-JUN-2001; 2001US-0295890P.
 PR 09-JUL-2001; 2001US-0303899P.
 PR 31-AUG-2001; 2001US-0316362P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM,
 PI WPI; 2003-093124/08.
 DR
 XX
 PT New purified or isolated nucleic acids of essential genes of Aspergillus
 PT fumigatus, useful for treating or preventing infections by A. fumigatus,
 PT or for treating a non-infectious disease in a subject e.g. cancer.
 XX
 PS Disclosure; Page; 175pp; English.
 XX
 CC The invention relates to novel purified or isolated nucleic acids of
 CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
 CC the invention are used to treat or prevent infections by a pathogenic
 CC organism such as A. fumigatus, to treat a non-infectious disease in a
 CC subject (e.g. cancer), to prevent or inhibit formation of an object
 CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
 CC biofilm comprising A. fumigatus. The polynucleotides are useful for
 CC expressing recombinant protein for characterization, screening or
 CC therapeutic use, as markers for host tissues in which the pathogenic
 CC organisms invade or reside, for comparing with the DNA sequence of A.
 CC fumigatus to identify duplicated genes or paralogues having the same or
 CC similar biochemical activity and/or function, for comparing with DNA
 CC sequences of other related or distant pathogenic organisms to identify
 CC potential orthologous essential or virulence genes, for selecting and
 CC making oligomers for attachment to a nucleic acid array for examination
 CC of expression patterns, for raising anti-protein antibodies, as an
 CC antigen to raise anti-DNA antibodies or to elicit another immune
 CC response, and for identifying polynucleotides encoding the other protein
 CC with which binding occurs or to identify inhibitors of the binding
 CC interaction. The polypeptides may be used to raise antibodies or to
 CC elicit immune response, as a reagent in assays designed to quantitatively
 CC determine levels of the protein in biological fluids, as a marker for
 CC host tissues in which pathogenic organism invade or reside, and to
 CC isolate correlative receptors or ligands in the case of virulence
 CC factors. This polynucleotide sequence represents one of the essential
 CC genes of Aspergillus fumigatus of the invention
 CC
 XX
 SQ Sequence 1410 BP; 354 A; 377 C; 350 G; 329 T; 0 U; 0 Other;
 Query Match 15.2%; Score 167.6; DB 8; Length 1410;
 Best Local Similarity 52.6%; Pred. No. 3.5e-41;
 Matches 418; Conservative 0; Mismatches 364; Indels 12; Gaps 2;
 QY 323 AAGAGTTTCACAGATGATGATGAGCTGTTAAAGTAAAGGACACAACTGTTCCATGA 382
 DB 617 AACGGTTTGGCAGAGCTGACATATCAACGCGACACGACGTCGTGACACCG 676
 QY 383 TGGCTCTGGGTGGAACCGAGTGAAGAAAGATGAATCTACGAAAAGCTTGATGAGA 442
 DB 677 TAGCCAGAGGCATTTTGAATGGAAGCGGAAAGCCAAAGTTTACAAATGATCCACCG 736
 QY 443 TTCATCAGTTTCTTGAATGCTTCTACTGTTGATAGGATGCGGTATGCTTATCGGCG 502
 DB 737 TCCAAATCGTTCTGATGCTTTATATGTCCTGATTTGTTATAGCAATGTTAAATCGGTC 796
 QY 503 AGCATTTGAGTTG-----CATATCCAAACCCACCACTTCCACAGAGTGGTTATCATAC 556

DB 797 AACATATGCGCTTGACGGAACAGACGATGTTGCCACCCGAACCTACGTGGTATCATCT 856
 QY 557 AACACAGATGCTCTCATGATGAGTGCAAGAAATGCTAGTAAAGTGAAGTGCATTT 616
 DB 857 GCAACAAAGCAAAATGTCGAGAGAGTGCCTCGAGGCCATTTGCAAAAGCTGTTGCTCT 916
 QY 617 GTTTCAGAGATATGCTTCT-----GTCGCGAGATTAACATATATGCGATCCAAAGTT 670
 DB 917 GTAGGACTATGATGATGCTTTTTCGAAGCACCCGAAGATTCAGCTCATCTGCAAGACGACC 976
 QY 671 TCACCTTTTCGATATGTTGCGAACCCATTTGATGATTTATGTTATGATGATGATGATGAT 730
 DB 977 TCACCTTCAATGATGATGCTTCCGCTGACCTTCAACATGCTCTTTGAGACTTTAAGAAAT 1036
 QY 731 CTCTCGTGTGTCGAAGCCGTTTGTGACTCTGATGAGGTTGACCCACCAATCCGTA 790
 DB 1037 CGCTGCTGCTGTTGGTGAGACCCAGCGGCGGACAAAGAACATTCCTGTGACAAAG 1096
 QY 791 TCATTGTCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 850
 DB 1097 TCATTGTCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1156
 QY 851 TACCGAAGACGGTCTCCTTAATATTCATCTTACCTTACGACGACGACGACGACGACGAC 910
 DB 1157 TACCCGTTCTTCATCCCGTTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1216
 QY 911 TTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 970
 DB 1217 CCATTCGATCCGATCCGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1276
 QY 971 ATGCTGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1030
 DB 1277 AGGATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1336
 QY 1031 CCATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1090
 DB 1337 GTATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1396
 QY 1091 AGCCTTGCCATGA 1104
 DB 1397 AGCCCTGCAATGA 1410
 RESULT 15
 ABR20506
 ID ABR20506 standard; DNA; 1625 BP.
 XX ABR20506;
 AC
 XX 16-APR-2003 (first entry)
 DT
 XX
 DE Aspergillus fumigatus essential gene #2864.
 XX
 KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
 KW cancer; contamination; biofilm; antibody; immune response; ds.
 OS Aspergillus fumigatus.
 XX
 PN WO200286090-A2.
 PD 31-OCT-2002.
 XX
 PF 23-APR-2002; 2002WO-US013142.
 PR 23-APR-2001; 2001US-0285697P.
 PR 27-APR-2001; 2001US-0287066P.
 PR 05-JUN-2001; 2001US-0295890P.
 PR 09-JUL-2001; 2001US-0303899P.
 PR 31-AUG-2001; 2001US-0316362P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM,

XX WPI; 2003-093124/08.
 XX
 PT New purified or isolated nucleic acids of essential genes of *Aspergillus*
 PT fumigatus, useful for treating or preventing infections by *A. fumigatus*,
 PT or for treating a non-infectious disease in a subject e.g. cancer.
 XX
 PS Disclosure; Page; 175pp; English.
 XX
 CC The invention relates to novel purified or isolated nucleic acids of
 CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of
 CC the invention are used to treat or prevent infections by a pathogenic
 CC organism such as *A. fumigatus*, to treat a non-infectious disease in a
 CC subject (e.g. cancer), to prevent or contain contamination of an object
 CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a
 CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for
 CC expressing recombinant protein for characterization, screening or
 CC therapeutic use, as markers for host tissues in which the pathogenic
 CC organisms invade or reside, for comparing with the DNA sequence of *A.*
 CC *fumigatus* to identify duplicated genes or paralogues having the same or
 CC similar biochemical activity and/or function, for comparing with DNA
 CC sequences of other related or distant pathogenic organisms to identify
 CC potential orthologous essential or virulence genes, for selecting and
 CC making oligomers for attachment to a nucleic acid array for examination
 CC of expression patterns, for raising anti-protein antibodies, as an
 CC antigen to raise anti-DNA antibodies or to elicit another immune
 CC response, and for identifying polynucleotides encoding the other protein
 CC with which binding occurs or to identify inhibitors of the binding
 CC interaction. The polypeptides may be used to raise antibodies or to
 CC elicit immune response, as a reagent in assays designed to quantitatively
 CC determine levels of the protein in biological fluids, as a marker for
 CC host tissues in which pathogenic organism invade or reside, and to
 CC isolate correlative receptors or ligands in the case of virulence
 CC factors. This polynucleotide sequence represents one of the essential
 CC genes of *Aspergillus fumigatus* of the invention
 XX
 XX Sequence 1625 BP; 398 A; 432 C; 401 G; 394 T; 0 U; 0 Other;

Query Match 13.81; Score 151.8; DB 8; Length 1625;
 Best Local Similarity 52.54; Pred. No. 3.3e-36;
 Matches 386; Conservative 0; Mismatches 337; Indels 12; Gaps 2;

QY 323 AAGAGTTCACACAGATGATCAAGCTGTTAAAGTAAGCACAACAACGTGTTCCATGA 382
 DB 765 AAGGTTTGGAAGAGCGCTCAGCATATCAAGCGGACACAGAGCGTGTGACACCG 824
 QY 383 TGGCTCTGGGTGAAACAGCTGAAGAAAGAAATGAATCTTACGAAAGCTTGATGAGA 442
 DB 825 TAGCCAGGGCATTTGGAATGGAAGCGAAACGCCAAGTTTACAAATGATTCACCG 884
 QY 443 TTGATGAGTTCTTGATCGCTTCTACTGTCGTGTAAGGATCCGATGCTTATGGGGC 502
 DB 885 TCCAAATCGTTCAGATCGCTTTTATATGTCGGTATGATGCAATGTATATCGGTC 944
 QY 503 AGCATGTTGAGTG-----CATATCCAAACCCACCACTTACACAGTGGGTTACATAC 556
 DB 945 AACACATCGCCCTGAGGAAAGACAGCATGTTGGCCACCGAACTACGTGATATCATCT 1004
 QY 557 ACACCAAGATGTTCTCTATGAGGTGGCAAGAAATGCTAGTGAAGATGCAAGTCAATTT 616
 DB 1005 GCACCAAGACAAATGTCGGAAGAGTCGCCCTCGAGGCCATTTGCAATGCTGTTTGTCT 1064
 QY 617 GTTTCAGAGATGATGTTCT-----GTCGGAGATTAACATATATGGGATCCAAAGTT 670
 DB 1065 GTAGAGACTATAGTGGCTTTTGAAGACACCGAAGATTACGCTCATCTGCAAGAGACACC 1124
 QY 671 TCACTTTTCCGATGTTCCGACCCATTGTCATTATGATGATGATGATGATGATGATGAT 730
 DB 1125 TCAACTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1184
 QY 731 CTCTCCGTGCTGTCGAAGAGCGGTTGTTATCTGTATAGGTTGACACCAATCCGTA 790
 DB 1185 CGCTGCTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1244

QY 791 TCATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 850
 DB 1245 TCATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1304
 QY 851 TACCGAAGAGCGGTCTCCCTAAATATTCATTTACCTCTACAGCACTGCAAGAAACCCAC 910
 DB 1305 TACCCGTTCTTCATCCCGTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1364
 QY 911 TTGAAGAATGTCGACTTGGAAACCGCTGATGTTCCCTGATGATGCTGTTATGCTT 970
 DB 1365 CCAATCTGATTCGGAATTCGATTAAGACGACTTCAAGCTCCATGCTGCTGCTGCTGCTGCT 1424
 QY 971 ATGCTGCTTATTAAGTCTGCTGATGCTGCTGATGCTGCTGATGCTGCTGATGCTGCTGAT 1030
 DB 1425 ACGGATTCGCAATCACTGCTGCTGATGCTGCTGATGCTGCTGATGCTGCTGATGCTGAT 1484
 QY 1031 CCATGGAAGGATTCG 1045
 DB 1485 GTATGGAAGGATTCG 1499

Search completed: April 12, 2005, 05:23:03
 Job time : 704 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2005, 23:08:37 ; Search time 227 Seconds
(without alignments)
7957.927 Million cell updates/sec

Title: US-10-642-531-1

Perfect score: 1104

Sequence: 1 atgcgcgtgaagaagcctag.....cgcagagccttgcacatga 1104

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/5A COMB. seq:*

2: /cgn2_6/prodata/1/ina/5B COMB. seq:*

3: /cgn2_6/prodata/1/ina/5A COMB. seq:*

4: /cgn2_6/prodata/1/ina/5B COMB. seq:*

5: /cgn2_6/prodata/1/ina/5A COMB. seq:*

6: /cgn2_6/prodata/1/ina/5B COMB. seq:*

7: /cgn2_6/prodata/1/ina/5A COMB. seq:*

8: /cgn2_6/prodata/1/ina/5B COMB. seq:*

9: /cgn2_6/prodata/1/ina/5A COMB. seq:*

10: /cgn2_6/prodata/1/ina/5B COMB. seq:*

11: /cgn2_6/prodata/1/ina/5A COMB. seq:*

12: /cgn2_6/prodata/1/ina/5B COMB. seq:*

13: /cgn2_6/prodata/1/ina/5A COMB. seq:*

14: /cgn2_6/prodata/1/ina/5B COMB. seq:*

15: /cgn2_6/prodata/1/ina/5A COMB. seq:*

16: /cgn2_6/prodata/1/ina/5B COMB. seq:*

17: /cgn2_6/prodata/1/ina/5A COMB. seq:*

18: /cgn2_6/prodata/1/ina/5B COMB. seq:*

19: /cgn2_6/prodata/1/ina/5A COMB. seq:*

20: /cgn2_6/prodata/1/ina/5B COMB. seq:*

21: /cgn2_6/prodata/1/ina/5A COMB. seq:*

22: /cgn2_6/prodata/1/ina/5B COMB. seq:*

23: /cgn2_6/prodata/1/ina/5A COMB. seq:*

24: /cgn2_6/prodata/1/ina/5B COMB. seq:*

25: /cgn2_6/prodata/1/ina/5A COMB. seq:*

26: /cgn2_6/prodata/1/ina/5B COMB. seq:*

27: /cgn2_6/prodata/1/ina/5A COMB. seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	873.6	79.1	1457	US-09-355-912A-1	Sequence 1, Appl
2	873.6	79.1	1457	US-10-202-428-1	Sequence 1, Appl
3	524.6	47.5	1602	US-09-333-423-3	Sequence 3, Appl
4	448.2	40.6	1332	US-09-333-423-1	Sequence 1, Appl
5	324.2	29.4	899	US-10-101-464A-965	Sequence 965, App
6	208.2	18.9	371	US-10-101-464A-464	Sequence 464, App
7	120.2	10.9	624	US-09-248-796A-4156	Sequence 4156, App
8	80.8	7.3	1601	US-09-949-016-5850	Sequence 5850, App
9	79	7.2	1866	US-08-878-989-13	Sequence 13, Appl
10	79	7.2	1866	US-09-272-796-13	Sequence 13, Appl
11	71.8	6.5	1797	US-09-949-016-4569	Sequence 4569, App
12	71.8	6.5	1797	US-09-949-016-4569	Sequence 4569, App
13	62.4	5.7	1422	US-09-949-016-3858	Sequence 3858, App
14	60.8	5.5	1422	US-09-949-016-377	Sequence 377, App
15	60.8	5.5	1422	US-09-949-016-17592	Sequence 17592, A
16	54.6	4.9	1864	US-09-949-016-12119	Sequence 12119, A
17	54.6	4.9	1864	US-09-949-016-15598	Sequence 15598, A
18	53.4	4.8	601	US-09-949-016-163093	Sequence 163093, A
19	53.4	4.8	601	US-09-949-016-163094	Sequence 163094, A
20	53.4	4.8	601	US-09-949-016-163094	Sequence 163094, A
21	45.4	4.1	738	US-09-248-796A-2138	Sequence 2138, App
22	39.6	3.6	1141	US-09-806-708B-22	Sequence 22, Appl
23	38	3.4	3176	US-08-425-299A-3	Sequence 3, Appl
24	38	3.4	3176	US-08-486-663A-14	Sequence 14, Appl
25	38	3.4	3176	US-08-247-804B-13	Sequence 13, Appl
26	38	3.4	3176	US-08-767-942A-14	Sequence 14, Appl
27	38	3.4	3176	US-09-070-060-10	Sequence 10, Appl

28	38	3.4	3177	US-09-357-746-10	Sequence 10, Appl
29	38	3.4	7218	US-08-232-463-14	Sequence 14, Appl
30	37	3.4	2817	US-09-949-016-1916	Sequence 1916, App
31	37	3.4	3063	US-08-294-312B-3	Sequence 3, Appl
32	37	3.4	3063	US-08-468-024B-3	Sequence 3, Appl
33	37	3.4	3063	US-09-708-200-12	Sequence 12, Appl
34	37	3.4	3063	US-09-023-655-1435	Sequence 1435, App
35	37	3.4	3063	US-08-465-679-3	Sequence 3, Appl
36	37	3.4	3063	US-09-788-657-8	Sequence 8, Appl
37	37	3.4	3063	US-09-712-691-10	Sequence 10, Appl
38	37	3.4	3063	US-09-707-468C-10	Sequence 10, Appl
39	36.6	3.3	154746	US-09-827-688-8	Sequence 8, Appl
40	34.8	3.2	231	US-09-248-796A-7339	Sequence 7339, App
41	34.8	3.2	3177	US-09-826-312A-1	Sequence 1, Appl
42	34.8	3.2	3177	US-09-542-497A-1	Sequence 1, Appl
43	34.4	3.1	431	US-09-513-999C-23703	Sequence 23703, A
44	34.2	3.1	2235	US-09-153-804-2	Sequence 2, Appl
45	34.2	3.1	3079	US-09-643-597-116	Sequence 116, App

ALIGNMENTS

RESULT 1									
US-09-355-912A-1									
; Sequence 1, Application US/09355912A									
; Patent No. 6500670									
; GENERAL INFORMATION:									
; APPLICANT: Zou, Jitao									
; TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE									
; FILE REFERENCE: 40942-1									
; CURRENT APPLICATION NUMBER: US/09/355,912A									
; PRIOR FILING DATE: 1999-10-15									
; PRIOR APPLICATION NUMBER: PCT/CA98/00096									
; PRIOR FILING DATE: 1998-02-09									
; PRIOR APPLICATION NUMBER: US 60/038,815									
; NUMBER OF SEQ ID NOS: 9									
; SOFTWARE: Patent version 3.0									
; SEQ ID NO 1									
; LENGTH: 1457									
; TYPE: DNA									
; ORGANISM: Arabidopsis thaliana									
US-09-355-912A-1									
Query Match									
Best Local Similarity 87.6%; Pred. No. 9.6e-293;									
Matches 967; Conservative 0; Mismatches 134; Indels 3; Gaps 1;									
QY	1	ATGCGCGTGAAGAGCTGCGAGTCTTGGAGAGCTTGAGAGCGTTCCAGCA	60						
DB	104	ATGCGCGTGAAGAGCTGCGAGTCTTGGAGAGCTTGAGAGCGTTCCAGCA	163						
QY	61	TGGGATGATGAAGAGCTGCGAGTCTTGGAGAGCTTGAGAGCGTTCCACT	120						
DB	164	TGGGATGATGAAGAGCTGCGAGTCTTGGAGAGCTTGAGAGCGTTCCACT	223						
QY	121	CCGCTGAGAGAACTTCTGATCTGGCGAGTTCTTCAAGAGCTTCCAGTTCCG	180						
DB	224	CCGCTGAGAGAACTTCTGATCTGGCGAGTTCTTCAAGAGCTTCCAGTTCCG	283						
QY	181	ATGCGCGTGAAGAGCTGCGAGTCTTGGAGAGCTTGAGAGCGTTCCAGCA	240						
DB	284	ATGCGCGTGAAGAGCTGCGAGTCTTGGAGAGCTTGAGAGCGTTCCAGCA	343						
QY	241	GTTTGAAGTGAAGAGCTGATGATGAGTCAATTCAGGAGCATGAGAGCTTCTCGAG	300						
DB	344	GTTTGAAGTGAAGAGCTGATGATGAGTCAATTCAGGAGCATGAGAGCTTCTCGAG	403						
QY	301	ATCAAGGATGCTGCTGATGAGAGAGTTTCAACAGATGATCAAGGCTTTAAAGTAAAG	360						
DB	404	ATCAAGGATGCTGCTGATGAGAGAGTTTCAACAGATGATCAAGGCTTTAAAGTAAAG	463						

```

QY 361 CACAACAGTGGTTCCTCCATGATGCTCTGGGTGACACCACTGAGAAAGATGAA 420
Db 464 CATACAGATGCTTCCATGATGCTTGGGTGATTAACACTCAGAAAGAAATGAA- 522
QY 421 CTCTACGAAAAGCTTGATGATGATTCATGATTCATTCGCTTCTTACTGCTGATPA 480
Db 523 --TTCTGGAATCTTGATGATGATTCATGATTCATTCGCTTCTTACTGCTGATPA 580
QY 481 GGGATCCGATGCTTATGCGGAGCATGTTGATGCTATTCATTCACCAACCACTTAC 540
Db 581 GGGATCCGATGCTTATGCGGAGCATGTTGATGCTATTCATTCACCAACCACTTAC 640
QY 541 ACAGTGGGTTATACACACCAAGATGCTCTATGAGAGTGGCAAGGATGCTGATGA 600
Db 641 ACAGTGGGTTATACACACCAAGATGCTCTATGAGAGTGGCAAGGATGCTGATGA 700
QY 601 GATGCAAGTGCATTTGTTTTCAGAGATGATGCTCTGCTCCGAGATTAACATATGCG 660
Db 701 GATGCTCGGTCAATTTGTTTTCAGAGATGATGCTCTGCTCCGAGATTAACATATGCG 760
QY 661 GATCCAGTTCATCTTCCGATGCTCCGATGCTCCGATGCTCCGATGCTCCGATGCT 720
Db 761 GATCCAGTTCATCTTCCGATGCTCCGATGCTCCGATGCTCCGATGCTCCGATGCT 820
QY 721 GTCAAGAACTCTCTCCGATGCTCCGATGCTCCGATGCTCCGATGCTCCGATGCT 780
Db 821 GTCAAGAACTCTCTCCGATGCTCCGATGCTCCGATGCTCCGATGCTCCGATGCT 880
QY 781 CCAATCCGATTCATGTTGCTGATGAGATTCGAAATGTTTAAAGTTCATGATGA 840
Db 881 CCAATCCGATTCATGTTGCTGATGAGATTCGAAATGTTTAAAGTTCATGATGA 940
QY 841 GGTGAGGTATACCGAGAGAGGCTCTCCCTAAATATTCATCTTACTCTACAGACATGA 900
Db 941 GGTGAGGTATACCGAGAGAGGCTCTCCCTAAATATTCATCTTACTCTACAGACATGA 1000
QY 901 AGAAACCCGCTTGGAGAGATGCTGAGATTCGAAATGCTGAGATTCGAAATGCT 960
Db 1001 AGAAACCCGCTTGGAGAGATGCTGAGATTCGAAATGCTGAGATTCGAAATGCT 1060
QY 961 GGTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 1061 GGTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1120
QY 1021 CAGATCATATTCATGAGAGATGCTGAGATTCGAAATGCTGAGATTCGAAATGCT 1080
Db 1121 CAGATCATATTCATGAGAGATGCTGAGATTCGAAATGCTGAGATTCGAAATGCT 1180
QY 1081 GACTCGCAGAGCCTTGGCCATGA 1104
Db 1181 GATTCCGCAAGAGCTTTACCTCTGA 1204

```

```

RESULT 2
US-10-202-428-1
; Sequence 1, Application US/10202428
; Patent No. 6825039
; GENERAL INFORMATION:
; APPLICANT: Zou, Jitao
; TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE
; FILE REFERENCE: 40942-1
; CURRENT APPLICATION NUMBER: US/10/202,428
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/09/355,912
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/CA98/00096
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/038,815
; PRIOR FILING DATE: 1997-02-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0

```

```

; SEQ ID NO 1
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-202-428-1

Query Match 79.1%; Score 873.6; DB 4; Length 1457;
Best Local Similarity 87.6%; Pred. No. 9,6e-293;
Matches 967; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

QY 1 ATGCGCGTGAAGAAAGCTGACGATGCTTTCGAAAGCTTGTATGAGACCTTACACAG 60
Db 104 ATGCGCGTGAAGAAAGCTGACGATGCTTTCGAAAGCTTGTATGAGACCTTACACAA 163
QY 61 TGGGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 164 TGGGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 223
QY 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGCTTCTTACAGAGAGCTTCCGATTCG 180
Db 224 CCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 283
QY 181 ATGCGAGAGCGTGCATGCACTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 284 ATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 343
QY 241 GTCTTGAAGTAAAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 300
Db 344 GTCTTGAAGTAAAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 403
QY 301 ATCAAGATATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 404 ATCAAGATATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 463
QY 361 CACAACAGTGGTTCCTCCATGATGCTGAGTGAACCAAGTGAAGAAAGATGAA 420
Db 464 CACAACAGTGGTTCCTCCATGATGCTGAGTGAACCAAGTGAAGAAAGATGAA- 522
QY 421 CTCTACGAAAAGCTTGATGATGATTCATGATTCATTCGCTTCTTACTGCTGATPA 480
Db 523 --TTCTGGAATCTTGATGATGATTCATGATTCATTCGCTTCTTACTGCTGATPA 580
QY 481 GGGATCCGATGCTTATGCGGAGCATGTTGATGCTATTCATTCACCAACCACTTAC 540
Db 581 GGGATCCGATGCTTATGCGGAGCATGTTGATGCTATTCATTCACCAACCACTTAC 640
QY 541 ACAGTGGGTTATACACACCAAGATGCTCTATGAGAGTGGCAAGGATGCTGATGA 600
Db 641 ACAGTGGGTTATACACACCAAGATGCTCTATGAGAGTGGCAAGGATGCTGATGA 700
QY 601 GATGCAAGTGCATTTGTTTTCAGAGATGATGCTCTGCTCCGAGATTAACATATGCG 660
Db 701 GATGCTCGGTCAATTTGTTTTCAGAGATGATGCTCTGCTCCGAGATTAACATATGCG 760
QY 721 GATCCAGTTCATCTTCCGATGCTCCGATGCTCCGATGCTCCGATGCTCCGATGCT 780
Db 821 GATCCAGTTCATCTTCCGATGCTCCGATGCTCCGATGCTCCGATGCTCCGATGCT 880
QY 841 CCAATCCGATTCATGTTGCTGATGAGATTCGAAATGTTTAAAGTTCATGATGA 900
Db 941 CCAATCCGATTCATGTTGCTGATGAGATTCGAAATGTTTAAAGTTCATGATGA 1000
QY 901 GGTGAGGTATACCGAGAGAGGCTCTCCCTAAATATTCATCTTACTCTACAGACATGA 960
Db 1001 GGTGAGGTATACCGAGAGAGGCTCTCCCTAAATATTCATCTTACTCTACAGACATGA 1060

```

```

QY      961  GGTAATGGTAATAGGCTGCGTAATAGTGGCTTGATAGTCGCTATTTTGGAGGATTTG 102
DB      1061  GGATATGGTATATGCTCTTCCTCCAAATTAGTGGCTTGATAGCTCCAAATATTTGGATGAGATTGG 1120
QY      1021  CAGATCATATCCATGGAAGGATATACGGGACTGATGCTTACTTGCACTTCTCTGTTGGGA 1080
DB      1121  CAGATCATATCCATGGAAGGATATGAGGACTGATGCTTACTTGCACTTCTCTGTTGGGA 1180
QY      1081  GACTGCGCAGAGGCTTTGGCCATGA 1104
DB      1181  GATTGCGAAGGCTTTTACCTCTGA 1204

RESULT 3
US-09-333-423-3
; Sequence 3, Application US/09333423
; Patent No. 6265636
; GENERAL INFORMATION:
; APPLICANT: Randall, Douglas
; APPLICANT: Thelen, Jay
; APPLICANT: Miernyk, Jan
; APPLICANT: Muzynski, Michael
; APPLICANT: Sewalt, Vincent
; TITLE OF INVENTION: Pyruvate Dehydrogenase Kinase
; FILE REFERENCE: 0818
; CURRENT APPLICATION NUMBER: US/09/333,423
; CURRENT FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,998
; EARLIER FILING DATE: 1998-06-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1602
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)...(1169)
US-09-333-423-3

Query Match      47.5%; Score 524.6; DB 3; Length 1602;
Best Local Similarity 68.8%; Pred. No. 3e-171;
Matches 760; Conservative 0; Mismatches 324; Indels 21; Gaps 2;

QY      12  GAAGGCTACGAGAGATGTTTTCGAAGAGCTTGATCGAGGACGTTCAAGATGAGGATGCAT 71
DB      77  GATGGCGTCGAGGCGCGGTGCGCGGGCGGTGGCGGAGGATGGGCGCTCGGGGACAGCAT 136
QY      72  GAAGGAGACGGGGCGGACCTTCAGTTAATGATGGAAGTTTCGCTTCCACTCCCACTGAGAG 131
DB      137  GAAGGAGACGGGGGAGACCTTCGCGGTACATGATGGAAGTTTCGCTTCCGCCACCCACAGCG 196
QY      132  AAACCTTGTGATCTTCGGCGGCAATTTCTTCAAGAGAGCTTCGCAATTCGATTCGCGAGGCG 191
DB      197  CAACCTGCTCTCTTCGGCGGCAATTCCTGCAAGAGAGCTCCCACTCGCCTTCGGCGCGCG 256
QY      192  TGCAGTGAACCTCGAGACGCTGCTCTTATAGCCCTCTTGAGAAACCTGCGCTTGAAGGT 251
DB      257  CGCGCTCGAGCTCGACTGCTGCTGCCCTTTCGGGCTCTTCAAGAGCCCGCATCTTCAAGGT 316
QY      252  AAGAGATTGTATGTGAGAGTCAATTCAGGAGCACTGAGACGTTTCTTGAGATCAAGAGATAC 311
DB      317  GCGGAGACTGTATCTTGAGACTCAATTCGGGACACTCAGATATCTTCCCTGAAGTGAAGAGCG 376
QY      312  TGCTGATGAGAAAGAGTTCAACAGATGATCAAGGCTGTTAAAGTAAGGACAAACAAGT 371
DB      377  GAAGGAGAGGCTCGCTTTCAGGCAATATATCAATATGATCAAGAGTGGCGGCAATTAACATGT 436
QY      372  GGTTTCCCATGATGGCTCTGGGGTGTGAACCAAGTGAAGAAAGAGATGAACCTTACGAAA 431
DB      437  GGTTTCCCAATATGGCTTGGAGGTGCACAGCTGTAAGAAAGAGCTGGGCGCCTCAAGGAA 496

```

```

QY      432  GCTT-----CATGAGATTCAATCAAGTTCTTGATAGGCTTCACTTGCTCGAT  479
Db      497  GGTTCATTCGAGTTCCAGTGGAGATTCACAGAGTTCCTTGAACGGTTCTTACATGTCAAGAT  556
QY      480  AGGATCCGTAATGCTTAATCCGGGACGAAATGTTGAGTTGCATTAATCCAAACCCACCACTTCA  539
Db      557  TGGATCCGATGCTGTAATAGGGACGAAATGTGGCTTTGCATGACCTTAAACCGAGGCTGG  616
QY      540  CACAGTGGGTTACATACACACCAAGATGTCTCTATGGAAGTGGCAAGAAATGCTAGTGA  599
Db      617  TGTATAGGCTCTCATCAATACMAATTAATCCCAATCAGGTGGCTCAAGCTGCTTGTGA  676
QY      600  AGATGCAAGGTCGATTTGGTTTCAGAGATATGTTCTGCTCCGGAGATPAAACATATATGG  659
Db      677  AGATGCGCGTTCTGTCTGTTTAAGGAAATATGATCAGCTCTCTATATTAACATTTATGG  736
QY      660  CGATCCAAAGTTTCACTTTTCCGTAATGTCGACACCATTGGATCTTAATGGTATGAGTT  719
Db      737  CGATCCAAATTTTACGTTTCATACGTACATTAACCTACACCTGATGTTATTTGAAAT  796
QY      720  AGTCAGAAATCTCTCTCCGTGTCTTCACAGACGGTTGTTGACTCTGATATAGGTTGCACC  779
Db      797  GGTAAAGAACTCTCTCCGCGCGGTCAGAAAGACGTAATATGAATCTGATATAAGATGTTCC  856
QY      780  ACCAATCCGTAATTTGTTGCTGATGGAATCGAAGATGTTACATTAAGGCTCAGATGA  839
Db      857  CCGCGTTAGAAATTTATAGTTGCTGATGGAAGAAAGATGTTATCTATCAAGGTTAGCATGA  916
QY      840  AGGTGAAGATATACCGAAGACCGGTCTTCCCTAAATAATCACTTACCTTACAGCACTGC  899
Db      917  AGGTGGCGGATACCAAGAAACGGGCTCTCCGAGAATTTTCAAGTATCTGTACACACACGC  976
QY      900  AAGAAACCCCACTTGAAGAAATGTGAGACTTGGAAACCGCTGATGTTCCCTGATTAATGGC  959
Db      977  AAGAAATCTCTCTAGCTAGTAAAT-----AGACCTAAATACTGAAGGTAACCTATAGGC  1027
QY      960  TGGTTATGTTATGATGTCGCTATTTAGTCGCTGTATAGTCTGCTAATTTTGGTGAATTT  1019
Db      1028  TGGGTAATGTTTTCGGGCTCCCGATCAGCCGCACTTACGCTCGATATTTTGGCGGCACT  1087
QY      1020  GCAGATCATATCATAGAAAGGATACGGGACTGATGCTTACTTGACACTTGTCTGCTTGG  1079
Db      1088  GCAATATCATCTCGATGGAAGGATACGGTACCGATGCTTACCTCACTTGTTCGCGGCTGG  1147
QY      1080  AGACTCGCAGAGGCTTGGCCATGA  1104
Db      1148  AGATTCAAGAGAGCCCTTGCCTTGA  1172

RESULT 4
US-09-333-423-1
; Sequence 1, Application US/09333423
; Patent No. 6265636
; GENERAL INFORMATION:
; APPLICANT: Randall, Douglas
; APPLICANT: Thelen, Jay
; APPLICANT: Muenzky, Jan
; APPLICANT: Mueszynski, Michael
; APPLICANT: Sewalt, Vincent
; TITLE OF INVENTION: Pyruvate Dehydrogenase Kinase
; TITLE OF INVENTION: Polynucleotides, Polypeptides and Uses Thereof
; FILE REFERENCE: 0818
; CURRENT APPLICATION NUMBER: US/09/333,423
; CURRENT FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,998
; EARLIER FILING DATE: 1998-06-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Zea mays

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (55) ... (1095)
US-09-333-423-1

```

Query Match	40.6%	Score 448.2;	DB 3;	Length 1332;
Best Local Similarity	66.1%	Pred. No. 1e-144;		
Matches 732; Conservative	0;	Mismatches 303;	Indels 72;	Gaps 3

QY	10	AAAGAGCTTACGAGAAAGTTTTCGAAAGACTTATCGAGAGCGTTACAGATGGGGATGC	69
Db	52	AGATAGCGCTCGAGCGCGGTGCGCGGCGGTGGCGAGAGGTGGCGGCTGGGGCGCG	111
QY	70	ATGAGCAGACGGGCGCTGAGCCCTCAGGTACATGATGGAGTTCCGCTTCCACTCCCATCTGAG	129
Db	112	ATCGCGCAGACCGGGGTCAAGCTTGGCGCTCACATGATGAGATTTGGGGCGGCGCCACAGAG	171
QY	130	AGAAACCTTCTGATCTCGGGCGCATTTCTTCAACAAGACTTCCGATTTGGATTCGCGAG	189
Db	172	CGGACCTTCTTCTCCCGCGCAGTTCTTGACACAAGAGCTCCCATTCGCATTCGCGGGA	231
QY	190	CGGCGCATCGAATCTGAGAGCGCGCCCTTAAGGCGCTCTCGAAGAAACCTGCGCTTTGAG	249
Db	232	CGGGCGCTGACCTCGACTCGCTCCCTTCGCGCTCTCCACCAACCGCCCATCTCTCAAG	291
QY	250	GTAAGAGATTGGTATGTGAGATCATTCACAGGACATGAGACGTTTCTGATGATCAAGAT	309
Db	292	GTCGAAGATTGGTATGTGAGTCAATTCGCTGAATCCGCTCTTTCAGAGGTGAGAAAC	351
QY	310	ACTGCTGATGAGAAAGATTCACACAGATGATCAAGCTGTTAAAGTAGGACACAACAC	369
Db	352	CAGAAAGATGACCTCGCTTCAACCAAGATGATCAAAATGATCAGGGTACGACACACAAAT	411
QY	370	GTCGTTCCATGATGGCTCTGGGTGGAACCAAGCTGAAGAAAG-----GATG	417
Db	412	GTGGTGCCCTGCGATTCATGGGGTGTGACGACCTGAAGAGACTTGGGCGGTCCAAAG	471
QY	418	AAACTCTACGAAAGCTTGATGAGATTCATCAGTTCTTGATTCGCTTACTTGTGTCGT	477
Db	472	GCGTTTCCCCCGGAATCATGATCATCATGATTCAGTTCTTGACCGCTTCAATGTCAAG	531
QY	478	ATAGGAGATCCGTAATGCTTATCGGGACGATGTGATGTCATTAATCCAAACCAACTT	537
Db	532	ATTGGGATCCGGATCTGATAGGGCTT-----	558
QY	538	CACACAGTGGGTATCATACACACCAAGATGTCCTCTATGAGAGTGGCAGAGAAATGCTAGT	597
Db	559	-----ATTAACACAAAATATGTCACCAAGACAGATGTCGAAATGCTAGT	603
QY	588	GAGATGCAAGGTGCAATTTGTTTTCAGAGATATGGTTCTGCTCCGAGATTAACATATAT	657
Db	604	GAAAGTACACGTGCTATTTGATGCGGGAGATATGGATCATCTCGATATGTCACATATAT	663
QY	658	GGCGATCCAAATTCATCTTTCCGTAATGTTCCGACCCATTTGCAATCTTAATGTGTATGAG	717
Db	664	GGTAGTCGGGTTTAACTTCATTCATATGTTTACACACATCTCAATCTCATGATTTTGAG	723
QY	718	TTAGTCAAGAACTCTCTCCGTGCTGCCAAGCGGTTGTTGACTCGATATAGGCTGCA	777
Db	724	CTGGTGAAGAACTCCCTTCGCGCGGTACAGAAACGATATGATTTCTGATATAGCTTGA	783
QY	778	CCACCAATCCGTAATCATTTGTTGCTGATGGAATCGAAGATGTTTACATTAAGAGTTCAGAT	837
Db	784	CCTCCAGTTCCAAATCATAGTTGCTGATGAGACAGAGATGTAATTAAGATTAGAT	843
QY	838	GAAGTGAGGTATACCGAGAAAGGGTCTCCCTAAATATTCATTACTTACCTTACAGCACT	897
Db	844	GAAAGTGTGAAATACCAAGAGTGGGCTTCAGAAATTTTTCATCTCTATATAGTACA	903
QY	898	GCAAGAAACCACTTGAAGAAGATGTGGAATTGGAACCGGTGATGTTCCCTGACATG	957
Db	904	GCAGAAAACCTCTGATCTGAGAGGACA-----TAAAGAGAGTAACTTATG	951

Qy	958	GCAGGTAATGGTTATAGCTCGCCCATATAGTCGTATAGTGGCTATTTTGGTGAAGAT	1017
Db	952	CTGGAGATATGGTTATGGGATCCCTTAATAGTCGCTTTATGCTGATATTTGGGTGGGAC	10111
Qy	1018	TTTGAGATCATATTCATGAGAAAGATACGGGACGTATGCTTACATTCGTAAGCTGTCGTCCTT	10777
Db	1012	TTTGAGATCATCTTATAGAAAGGATATGAAACGTATCCTTACCTCAACCTTTCACGGGCTG	10712
Qy	1078	GGAGACTCGAGAGCCCTTTGCCATGA	1104
Db	1072	GGAGATTCGAGGAAACCTTGCCCTAA	1098

RESULT 5
US-10-10

```

/ US-10-101-464A-965
/ Sequence 965, Application US/10101464A
/ Patent No. 6768041
/
/ GENERAL INFORMATION:
/ APPLICANT: Scribala, Timothy
/ APPLICANT: Nieuwenhuizen, Nicolaas
/ APPLICANT: Higgins, Colleen M.
/ TITLE OF INVENTION: Compositions Isolated from Plant Cells
/ TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
/ FILE REFERENCE: 11000.1020c2
/
/ CURRENT APPLICATION NUMBER: US/10/101.464A
/ CURRENT FILING DATE: 2002-03-18
/ PRIOR APPLICATION NUMBER: 09/704,302
/ PRIOR FILING DATE: 2000-11-01
/ PRIOR APPLICATION NUMBER: 09/228,986
/ PRIOR FILING DATE: 1999-01-12
/ PRIOR APPLICATION NUMBER: 60/162,866
/ PRIOR FILING DATE: 1999-11-01
/ PRIOR APPLICATION NUMBER: PC/US00/00724
/ PRIOR FILING DATE: 2000-01-11
/
/ NUMBER OF SEQ ID NOS: 89
/
/ SOFTWARE: FastSeq for Windows Version 4.0
/
/ SEQ ID NO 965
/
/ LENGTH: 899
/
/ TYPE: DNA
/
/ ORGANISM: Eucalyptus grandis
/
/ US-10-101-464A-965

```

Query Match	29.4%	Score 324.2;	DB 4;	Length 899;
Best Local Similarity	76.0%	Pred. No. 11e-101;		
Matches 414;	Conservative 0;	Mismatches 128;	Indels 3;	Gaps 1

QY	560	CCAGATGTCCTCTATGAGGGTGGCCAGAAATGCTAGTGAATGCAAGTCGATTTGTT	619
Db	144	CGAAGATGTCCTCCAGTTGAGGTTGCCAGTATGCTATGAGATGCTGTGCTATTTGTT	203
QY	620	TCAGAGATAGCTTCCTGCTCCGAGATTAACATATATGGGATTCGAAGTTTCACTTTTC	679
Db	204	TGCGGAGATAGGAGATGCTCTCGGATTTAATATATATGGGGATCCCAAGCTTTTCAATCC	263
QY	680	CGATGTTCCGACCCATTTGATGCTTATATGGTATGATTTAGTCAAGAATCTCTCCGCTG	739
Db	264	CGATGTTCCAAACACACTGTGATCTTATATGATATTTGACTTGTTAAGACCTCTTTGCTG	323
QY	740	CTGTCCAAAGACGGTTTGTGATCTCTGATAGGGTTGACCAACAACTCGATCATTTGTTG	799
Db	324	CTGTCCAAAGAACGATTCATGATTCAGATTAAGTTGACCTCTTATCGAATTAATTTGTTG	383
QY	800	CTGATGGAATCGAAGATGTTACAAATAAAGCTTCAGATGGAAGGTGANGATTAACCGAAG	859
Db	384	CTGATGAGACGAGACGTTACATCAAGGCTTCAGATGAGGGGGTGGCATTCCAAGAA	443
QY	860	GCGGTCTCCCTAAATATTCATTTACCTCTACAGACATGCAAGAAACCCACTTGAAAG	919
Db	444	GTGGTCTTCCCAAAATCTTACGATATGTGATATGACATCGGAAAAACCCCTGATATGA	503
QY	920	ATGTGACTTGGGAAACCGCTGATGTTCCCTGACTATAGCTGTGTTATGTTATGATGCTGC	979
Db	504	ACTCAGATCTTGGAAATAGTGAATAT---GTGACTATAGCCGGTATGAGGATATGGGCTTC	560

Db 1235 ATCTGACCGGGTCATGACTACCACTTCACTACTGCG 1271

RESULT 10

US-09-272-796-13

/ Sequence 13, Application US/09272796
/ Patent No. 6207148

GENERAL INFORMATION:

/ APPLICANT: Bandman, Olga

/ APPLICANT: Hillman, Jennifer L.

/ APPLICANT: Corley, Neil C.

/ APPLICANT: Guegler, Karl G.

/ APPLICANT: Lal, Preeti

/ APPLICANT: Goli, Surya K.

/ APPLICANT: Shah, Puri

/ TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN

/ TITLE OF INVENTION: KINASES

/ NUMBER OF SEQUENCES: 21

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Incyte Pharmaceuticals, Inc.

/ STREET: 3174 Porter Drive

/ CITY: Palo Alto

/ STATE: CA

/ COUNTRY: USA

/ ZIP: 94304

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Diskette

/ OPERATING SYSTEM: DOS

/ SOFTWARE: FastSeq for Windows Version 2.0

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/09/272,796

/ FILING DATE:

/ CLASSIFICATION:

/ PRIORITY APPLICATION DATA:

/ APPLICATION NUMBER: 08/878,989

/ FILING DATE:

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Billings, Lucy J.

/ REGISTRATION NUMBER: 36,749

/ REFERENCE/DOCKET NUMBER: PF-0321 US

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: 415-855-0555

/ TELEFAX: 415-845-4166

/ TELEX:

/ INFORMATION FOR SEQ ID NO: 13:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 1866 base pairs

/ TYPE: nucleic acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ IMMEDIATE SOURCE:

/ LIBRARY: COLTUT03

/ CLONE: 1340712

/ US-09-272-796-13

Query Match 7.2%; Score 79; DB 3; Length 1866;

Best Local Similarity 45.5%; Pred. No. 4e-16;

Matches 372; Conservative 0; Mismatches 430; Indels 15; Gaps 2;

Db 83 GCGTGAAGCTCAGTACATGATGAGTTCGATCCCACTGAGAGAACTTCTGA 142

470 GCTTACGCGCCACCATGATGCTCTACGCTGCGCTCTGAGAGAGGAGCATCTTCTGA 529

143 TCTCGGCGCATTTCTTTCACAGAGAGCTTCCATTTGGATCGGAGCGCTGCATGAAC 202

530 AAAGTGTCTGGTACCTGACAGCAAACTTCCAGTGAAGATTGCTCACCGCATCAAGGGCT 589

203 TCGAGACGCTGCTTATGAGCTCTCTGAGAACTGCGGCTTGAAGATTAAGATTGGT 262

590 TCGCGTCTGCTTCTTCTTCAATCATTTGGCTGCAACCCACCATCTGACGCTGATGAGCTAT 649

263 ATGTGAGTCAATTCAGGAGCATGAGAGCGTTTCTGATCAAGATTAATCTGATGAGA 322

Db 650 ATATCCGTCCTTCCAGAACTGACAGACTTCCCTCCGATCAAGAGACAGGCGGAGAG 709

323 AAGATTCACACAGATGATCAAGAGCTGTTAAAGTAAAGACCAACACCTGTTCCATGA 382

710 CCCAGTACGCGCAGCTGGTGCCAGACCTCTGATGACCAAGAGATGGTACCTCT 769

383 TGGCTCTGGGTGAAACAGCTGAAGAAAGAAATGAACTCTCAAGAAAGCTTGATGAGA 442

770 TGGCAGAGG-----CTTACGTGAAGCCGGAAGACATAGAGATGAAAGCTCG 820

443 TTCTACGATTTCTTGATCGCTTCTACTTGTCTGATTAAGGATCGGATGCTTATCGGC 502

821 TCGCTACTTCTTGACAAAGCTGACTGACGCTTGAAGCTTGAATCCGATGTTGGCCAGCG 880

503 AGCATTTAGTTGCTTAATCCAAACCCACACTTACACAGTGGTTTATACACACCA 562

881 ATCACCTGCGCTGATGAGACAGC-----TGACTTTGTGGCATATCTGTACTC 934

563 AGATGCTCTCTTATGAGAGTGGCAAGGATGCTAGTAAAGTCAAGATGATGATTTGTTCA 622

935 GTCTCTACCAAGAAATTTAGAGAGTGGGTGACTTTGCAACCGCTGTGAGC 994

623 GAGATGATGCTTCTGCTCCGAGATTAACATATATGCGCATCCAGTTTCACTTTCCGT 682

995 ACAAGTATGCAATGCGCCCGCTGTCGATCAATGCGCATGCTGCGCTCCCT 1054

683 ATGTTCCGACCATTTGATCTTATGATGATGATGATGATGATGATGATGATGATGAT 742

1055 TCAATCCCTATGCACTGAGCTACATCTCTCGAGTGTCTCAAGATGCAATGAGAGCA 1114

743 TCCAGAGCGGTTTGTGACTGCTGATAGGTTGACACCAATCCCTATCTATTTGCTG 802

1115 CAATGAGAGCCACCTAGACATCTCCCTACAATGTCCCAATGTGCTATCCATCCATGCGCA 1174

803 ATGAATGCAAGATTTACATTAAGGTTCTGATGAGTGGAGTATACGAGAGCG 862

1175 ACAATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1234

Db 863 GTCTCCCTAAATATCTACTTACCTTACAGACATGCG 899

1235 ATCTGACCGGGTCATGACTACCACTTCACTACTGCG 1271

RESULT 11

US-09-949-016-4569

/ Sequence 4569, Application US/09949016

/ Patent No. 6812339

/ GENERAL INFORMATION:

/ APPLICANT: VENTER, J. Craig et al.

/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

/ FILE REFERENCE: CL001307

/ CURRENT APPLICATION NUMBER: US/09/949,016

/ CURRENT FILING DATE: 2000-04-14

/ PRIOR APPLICATION NUMBER: 60/241,755

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-03

/ PRIOR APPLICATION NUMBER: 60/237,768

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR APPLICATION NUMBER: 60/231,498

/ NUMBER OF SEQ ID NOS: 207012

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 4569

/ LENGTH: 1797

/ TYPE: DNA

/ ORGANISM: Human

/ US-09-949-016-4569

Query Match 6.5%; Score 71.8; DB 4; Length 1797;

Best Local Similarity 47.6%; Pred. No. 1.2e-13;

Matches 451; Conservative 0; Mismatches 442; Indels 54; Gaps 6;


```

      808 CCTAAGCTGATGATGGAGAGGATGTCAGATGCTTGTAGTGTCAAGATGCTGT 867
      619 TTGAGAGAGATGTTCTGCTCCGGAGATAA-----CATATAGC 660
      868 GATCAGTATATTTATTCATCTCAAGATTAAGCTTACACAGTAAGTGAATTTCCA 927
      661 GATCAGATTTCACTTTCCGATGTTCCGACCATTTGATCTTATGTAAGTTA 720
      928 GACCAACATTTACATCGTATGTTCTTCTCACTCATATGCTCTTTGAAGTA 987
      721 GTCAAGAACTCTCTCCGTGCTGTCAGAGCGGTTGTTGACTGTGATAGGTTGACA 780
      988 TTTCAGAAATGCAATGCGGGCAGACATGTAACACAGGAAATCAGCTTCCCTT---ACA 1044
      781 CCAATCGATTCATTTGTTGCTGATGTAAGATGTTACATTAAGTCTCAGATGTA 840
      1045 CCAATAGAGTATATGTTGTTGTTGGGAAAGAAAGACCTTACATTAAGATTTCAAGACA 1104
      841 GGTGAGATATACCGAAGCGGCTCTCCATAATTCACCTTACCTCAGACACTGCA 900
      1105 GAGAGTGTGTTCTCCCTGAGAAATTTATGACCGCTCTTATGTTATACATCTCAGCTGCA 1164
      901 AGAAACCACTTGAAGAGATGTCAGCTGGAAACCGCTGATGTTCCCTGACTATGCT 960
      1165 CCAACCGCTGTGATGATTAATCCCGAATG-----CTCCTTGGCT 1206
      961 GGTATAGTTATGCTGCTTATGCTTATGCTGCTTATGCTGCTTATGCTGCTGCTGCT 1020
      1207 GGTGTTGTTACGGCTTGGCAATTTCTGCTGTATGCAAGATCTTCAAGAGATCTG 1266
      1021 CAGATCATATCCATGGAAGATACGGGATGATGCTTACTTGCATTT 1067
      1267 AATCTTACTCTTATTCAGATATGGAACAGATGCTATCTTACTT 1313

```

```

RESULT 13
US-09-949-016-3856
; Sequence 3856, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3856
; LENGTH: 1422
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3856

```

```

Query Match      5.7%; Score 62.4; DB 4; Length 1422;
Best Local Similarity 48.5%; Pred. No. 2e-10;
Matches 458; Conservative 0; Mismatches 426; Indels 60; Gaps 8;

```

```

      274 TTCAGGAGATGAGAGACGTTTCTCTGAGA-----TCAAGATATCTGATGAGAAAG 337
      323 CTCCTGACATCATATGAGATTTCTGACAGAGATCCGAGGACATCGACCTTGAGCAG 382
      328 TTACACAGATGATCAAGGCTGTTAAAGTAAAGCACAACAAGTGTGTTCCATGATGCT 387
      383 TTCACTGAGCGCCCTGTGTACCAATCCGAAACCGGACACAAGACGTGTGCTCCACATGCA 442
      388 CTGGGTGTGAACAGCTGGAAGAAAGAAAT---GAACTGTAGAAAGCTTGTATGATT 444
      443 CAAGCGCTGTGATGATCAAGAGACACTACGGGATGACCCGCTTCAACACGAAATC 502
      445 CATCAGTTCTTGATGCTTCTTATCTGCTGCTGATGAGATGAGATGCTTATCGGAG 504
      503 CAGTACTTCTGAGACCGCTTCTTACCTCAGCCGATCTCCATCCGATGCTCATCAACAG 562
      505 -----CATGTTAGTTGATATTCACAAACCCACACTTCACACAG--TGGGTACATA 555
      563 CACACCTCATCTTATGATGAGCAGCAACCAACCCGATCCCAACACATTCGGCAGACTC 622
      556 CACACCAAGATGCTCTATGAGAGTGGCAAGGATGCTAGTGAAGATGCAAGTGCATT 615
      623 GACCCCACTGCAACGCTCTGAGTGTGTCAGAAAGATGCTTACGACATGCTTACCTCTG 682
      616 TGTTCAGAGATGATGCTGCTCCGAGATTAACATATATG----- 658
      683 TGTGACAGATATTAATATGAGCTCCTGACCTGACCTGAGATCCAGAGATCAATGACGCAAC 742
      659 -GCGATCAAGTTTACCTTTCCGATGTTTCCGACCATTTGATCTTATGATGATGAG 717
      743 TCCAAACACCGATTCACATGATGCTACGCTCCCTCCACCTTACACATGCTCTTTGAG 802
      718 TTAGTGAAGATCTCTCCGCTGCTGTCGAAGAGCGGTTTGACTGTGATAGGTTGCA 777
      803 CTCTTAAGAAATGTCATGAGGCGGCTGTGGAAG---CCATGATTCACCTCATATCTC 859
      778 CCACCAATCCGATCATATGTTGCTGATGAGATGAGATGTTACATTAAGATCTCAGAT 837
      860 CCACCATCAAGGATCATGATGAGGCTTGGGATGAGAAATCTGTCATCAAGATGAGTAC 919
      838 GAAAGTGAAGTATACCGAAGCGGCTCTCCATAATTTCACTTACCTTACAGACT 897
      920 CGAGGTGGGGGTGTTCCCTTGAAGAAATGAGCGCTTCAAGCTACATGATCTCACA 979
      898 GCAAGAAACCACTTGAAGAGATGTCAGCTGGGAACCGTATGTTCCCTGACTATG 957
      980 GACCCACCCCTC-----AGCTGGCACCGGGGAAAGCGCC-----TG 1018
      958 GCTGTTATGATGATGCTGCTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
      1019 GCTGCTTGTGTTATGAGGCTCCCATTTCCCGCTCTACGCCAAGTACTTCCAGGAGAC 1078
      1018 TTGCAATCATATCCATGGAAGATACGGGATGATGCTTACTT 1061
      1079 CTGCACTCTTCTCATGGAAGGCTTTGGACCGAGCTGTGAT 1122

```

```

RESULT 14
US-09-949-016-377
; Sequence 377, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```

; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 377
 ; LENGTH: 1422
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-377

Query Match 5.5%; Score 60.8; DB 4; Length 1422;
 Best Local Similarity 48.4%; Pred. No. 7.1e-10;
 Matches 457; Conservative 0; Mismatches 427; Indels 60; Gaps 8;

```

QY 154 TTTCTTCAAGAGCTTCCGATTCGATTCGAGCCGCGCATGGAAGCTGAGACGGCTG 213
DB 202 TTCCTCAGGACGAGGTGCTGTGGCGCTGCGCCAAATCATGAAAGATCAAGCTGCTT 261
QY 214 CCTTATGCGCTCTCTGAGAAACCTGCGCTTTGAAAGTAAAGATTTGATGTGAGTCA 273
DB 262 CCCGACGAGTGTGAGACACCCAGTGCAGTGTGAGAGCTGTATGTCCAGAGC 321
QY 274 TTCAGGACATGAGACGTTTCTGAGA-----TCAGGATACGCTGATGAGAAAG 327
DB 322 CTCCTGACATCATGAGATTTCTTGACAGAGATCTTGAGGACCATGCACTTGAGCCG 381
QY 328 TTCACACAGATGATCAAGCTGTAAAGTAAAGGACCAACAGTGTTCCTGATGAGCT 387
DB 382 TTCACGACGCTGTGATCAATCCGAAACGGGACCAACGATGTGCGCCACATGGCA 441
QY 388 CTGGGTGAAACGAGCTGAAGAAAGAT---GAACTCTAGGAAAGCTTGATGAGATT 444
DB 442 CAGGCGGTCTGAGTACAGAGACCTACGCGCATGACCCGCTCCAAACGAGATC 501
QY 445 CATCAGTTCTTGAATCGCTTCTACTTGTCTGTATAGGATCCGATATGATGAGGAG 504
DB 502 CAGTACTTCTGAGACCGCTTCTACCTGACCGGCAATCCATCCGATGCTCATACAG 561
QY 505 -----CATGTTGAGTTGATATATCAAAACCAACCACTTACACAG-TGGTTACATA 555
DB 562 CACACCTGATCTTGAATGAGCAGACCAACCAACCCATCCCAAAACATGGGAGATC 621
QY 556 CACACCAAGATGTCTCTTATGAGTGAAGGAAAGTAAAGTAAAGTGAAGTGAAGTGA 615
DB 622 GACCCCAACTGCAACGCTCTGAGTGTGTAAGATCTTACGACATGAGTCTGAG 681
QY 616 TGTTCAGAGATGATGCTTCTGCTCGAGATTAACATATATG-----658
DB 682 TGTGACAGATGATGAGCTTCACTGACCTGAGATTCAGGAGATCAATGAGCCAC 741
QY 659 -GGATCCAGTTCACTTTTCCGATGTTCCGACCATTTGATGATGATGATGAG 717
DB 742 TCCAAACAGCGGATTCATGATGCTTACCTCCCTCCACCTTACCAACATGCTTTGAG 801
QY 718 TTAGTCAAGAACTCTCCGCTGCTGTCAGAGACGCTTTGATGATGATGAGTGA 777
DB 802 CTCCTTCAAGATGCAATGAGGAGGAGCTGTGGAAG---CCATGATCCAGCTCATTTTC 858
QY 778 CCACCAATCCGATCATGTTGCTGATGATGAAAGTAAAGTAAAGTCAAGTCAAGT 837
DB 859 CACCCCATCAAGGATGATGAGCTTGGGAGAGATCTGCTCAATCAAGATGAGTGA 918
QY 838 GAAGTGAAGTATACCGAAGACGGTCTCCAAATATTCCTTAACTTACCTTACAGACT 897
DB 919 CGAGTGGGGGTGTCTTCTTGAAGAAATTTAGAGCACTTTCAGCTTATATGACTTCA 978
QY 898 GCAGAAACCACTTGAAGAAATGAGCTTGGAAACCGCTGATGTTCCCTGACTATG 957
DB 979 GCAACCAACCCCTC-----AGCTGCAACGGGAGAGCCGCC-----TG 1017
QY 958 GCTGTTATGTTATGATGCTGCTTATGATGCTGCTGCTGCTGCTGCTGCTGCTG 1017
DB 1018 GCTGCTTGTGTTATGAGGCTCCCAATTTCCGCTTCTTACGCAAGTACTTCCAGGAGAC 1077
  
```

```

QY 1018 TTGCAATCATATTCATGAGAGATACGGGACTGATGCTTACTT 1061
DB 1078 CTGACGCTTCTTCCATGAGAAAGCTTTGGACCGAGCTGTGAT 1121
  
```

RESULT 15

```

US-09-949-016-17592
; Sequence 17592, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17592
; LENGTH: 72992
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-17592
  
```

Query Match 5.5%; Score 60.8; DB 4; Length 72992;
 Best Local Similarity 71.4%; Pred. No. 1.2e-08;
 Matches 80; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```

QY 957 GCTGTTATGTTATGATGCTGCTTATGATGCTGCTGCTGCTGCTGCTGCTG 1016
DB 68219 GCTGATTTGTTATGTTGTTGCTTCCGCTGTATGCTGATATTTTCAAGAGA 68278
QY 1017 TTTGAGATCATATTCATGAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1068
DB 68279 TCTGAACATGATTCATGAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 68330
  
```

Search completed: April 12, 2005, 05:29:23
 Job time : 317 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2005, 05:36:03 ; Search time 1555 Seconds
(without alignments)
4302.408 Million cell updates/sec

Title: US-10-642-531-1

Perfect score: 1 atgcgcgtgaagaagcctgagccttcgcacga 1104

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 5615251 seqs, 3030001701 residues

Total number of hits satisfying chosen parameters: 11230502

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09A_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10F_NEW_PUB.seq:
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1104	100.0	1104	17	US-10-222-075-1
2	1104	100.0	1104	18	US-10-642-531-1
3	1095.6	99.2	1104	17	US-10-222-075-2
4	1095.6	99.2	1104	18	US-10-642-531-2
5	1062.4	96.2	1104	17	US-10-222-075-4
6	1062.4	96.2	1104	18	US-10-642-531-4
7	1062	96.2	1104	17	US-10-222-075-3
8	1062	96.2	1104	18	US-10-642-531-3
9	873.6	79.1	1457	14	US-10-202-428-1
10	665.8	60.3	1382	13	US-10-062-254-13
11	559.4	50.7	2028	17	US-10-424-599-109588

12	547.4	49.6	1481	13	US-10-062-254-9	Sequence 9, Appl1
13	547.4	49.6	2892	18	US-10-437-963-74020	Sequence 74020, A
14	536.6	48.6	1165	17	US-10-425-114-15131	Sequence 15131, A
15	527.8	47.8	1612	17	US-10-425-114-15532	Sequence 15532, A
16	527.8	47.8	2130	18	US-10-425-115-151648	Sequence 151648, A
17	524.8	47.5	1424	13	US-10-062-254-23	Sequence 23, Appl1
18	524.6	47.5	1467	17	US-10-425-114-33770	Sequence 32770, A
19	523.6	47.4	1876	18	US-10-767-701-14658	Sequence 14658, A
20	522.2	47.3	1388	17	US-10-425-114-34871	Sequence 34871, A
21	522.2	47.3	1469	18	US-10-425-115-94621	Sequence 94621, A
22	521.4	47.2	1467	18	US-10-425-115-151647	Sequence 151647, A
23	512	46.4	1910	18	US-10-437-963-17757	Sequence 17757, A
24	511.6	46.3	1407	13	US-10-062-254-3	Sequence 3, Appl1
25	511	46.3	1502	17	US-10-425-114-4107	Sequence 4107, Ap
26	441.6	40.0	1100	18	US-10-437-963-74022	Sequence 74022, A
27	419.6	38.0	1394	17	US-10-425-114-34658	Sequence 34658, A
28	367.8	33.3	1114	13	US-10-062-254-21	Sequence 21, Appl
29	330.4	29.9	789	17	US-10-425-114-8176	Sequence 8176, Ap
30	328.2	29.7	591	18	US-10-021-332-9239	Sequence 9239, Ap
31	324.2	29.4	899	14	US-10-101-464A-965	Sequence 965, App
32	324.2	29.4	899	19	US-10-864-252-965	Sequence 2039, Ap
33	298.6	27.0	1378	18	US-10-424-599-110303	Sequence 110303, A
34	274.6	24.9	671	17	US-10-424-599-110303	Sequence 64166, A
35	274.4	24.9	2484	18	US-10-425-115-94622	Sequence 94622, A
36	266	24.1	2665	18	US-10-425-115-94622	Sequence 40472, A
37	263.8	23.9	590	17	US-10-424-599-109592	Sequence 109592, A
38	253	22.9	734	17	US-10-424-599-109592	Sequence 14799, A
39	253	22.8	734	17	US-10-425-114-14799	Sequence 7, Appl1
40	251.2	22.8	616	13	US-10-062-254-7	Sequence 1, Appl1
41	244.2	22.1	584	17	US-10-620-796-1	Sequence 11, Appl
42	233.4	21.1	503	13	US-10-062-254-11	Sequence 110305, A
43	232	21.0	741	17	US-10-424-599-110305	Sequence 1, Appl1
44	219.6	19.9	665	13	US-10-062-254-1	Sequence 16529, A
45	218.4	19.8	704	17	US-10-425-114-16529	

ALIGNMENTS

RESULT 1
US-10-222-075-1
; Sequence 1, Application US/10222075
; Publication No. US20040033606A1
; GENERAL INFORMATION:
; APPLICANT: Zou, Jitao
; TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE
; FILE REFERENCE: 3015-544205
; CURRENT APPLICATION NUMBER: US/10/222,075
; CURRENT FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: PDKK cDNA from B. napus
US-10-222-075-1

Query Match	100.0%	Score 1104;	DB 17;	Length 1104;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1104;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGCGGTGAAGAGCTGACGAGATGTTTTCGAAGAGCTTATCGAGACCTTCACAG	60	
DB	1	ATGGCGGTGAAGAGCTGACGAGATGTTTTCGAAGAGCTTATCGAGACCTTCACAG	60	
QY	61	TGGGATGATGAG	120	
DB	61	TGGGATGATGAG	120	
QY	121	CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCACAGAGAGCTTCGATTCCG	180	

```

Db      121  |||||
Qy      181  |||||
Db      181  |||||
Qy      241  |||||
Db      241  |||||
Qy      301  |||||
Db      301  |||||
Qy      361  |||||
Db      361  |||||
Qy      421  |||||
Db      421  |||||
Qy      481  |||||
Db      481  |||||
Qy      541  |||||
Db      541  |||||
Qy      601  |||||
Db      601  |||||
Qy      661  |||||
Db      661  |||||
Qy      721  |||||
Db      721  |||||
Qy      781  |||||
Db      781  |||||
Qy      841  |||||
Db      841  |||||
Qy      901  |||||
Db      901  |||||
Qy      961  |||||
Db      961  |||||
Qy      1021  |||||
Db      1021  |||||
Qy      1081  |||||
Db      1081  |||||

```

RESULT 2
US-10-642-531-1

; Sequence 1, Application US/10642531

; Publication No. US20040248302A1

; GENERAL INFORMATION:

```

; APPLICANT: Marillia, Elizabeth-France
; APPLICANT: Zou, Jitao
; TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE
; FILE REFERENCE: 3015-5442US
; CURRENT APPLICATION NUMBER: US/10/642,531
; PRIORITY FILING DATE: 2003-08-15
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; OTHER INFORMATION: PDHK cDNA from B. napus
US-10-642-531-1

Query Match      100.0%; Score 1104; DB 18; Length 1104;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  |||||
Db      1  |||||
Qy      61  |||||
Db      61  |||||
Qy      121  |||||
Db      121  |||||
Qy      181  |||||
Db      181  |||||
Qy      241  |||||
Db      241  |||||
Qy      301  |||||
Db      301  |||||
Qy      361  |||||
Db      361  |||||
Qy      421  |||||
Db      421  |||||
Qy      481  |||||
Db      481  |||||
Qy      541  |||||
Db      541  |||||
Qy      601  |||||
Db      601  |||||
Qy      661  |||||
Db      661  |||||
Qy      721  |||||

```


Db 721 GTCAGAACTCTCTCCGCTGTCCTCAAGAGCGGTTTGTGACTGATAGGGTTGCACA 780
Qy 781 CCAATCCGTAATCATGTTGCTGATGGAATGGAAGATGTTTAAATTAAGTCTCAGATGAA 840
Db 781 CCAATCCGTAATCATGTTGCTGATGGAATGGAAGATGTTTAAATTAAGTCTCAGATGAA 840
Qy 841 GGTGAGGTAATACCGAAGAGGGTCTCCCTAAATATTCCTTAACCTTAACAGACTGCA 900
Db 841 GGTGAGGTAATACCGAAGAGGGTCTCCCTAAATATTCCTTAACCTTAACAGACTGCA 900
Qy 901 AGAAACCCACTTGAAGAAGATGTGACTTGGGAAACCGGTATGTTCCCTGACTATGAGCT 960
Db 901 AGAAACCCACTTGAAGAAGATGTGACTTGGGAAACCGGTATGTTCCCTGACTATGAGCT 960
Qy 961 GGTATGATGATGCTGCTGCTATGATGCTGATGCTGCTATGCTGCTATGCTGCTGCTG 1020
Db 961 GGTATGATGATGCTGCTGCTATGATGCTGATGCTGCTGCTATGCTGCTGCTGCTGCTG 1020
Qy 1021 CAGATCATATCCATGGAAGATACGGGACTGATGCTTACTTGCATTTGCTGCTTGA 1080
Db 1021 CAGATCATATCCATGGAAGATACGGGACTGATGCTTACTTGCATTTGCTGCTTGA 1080
Qy 1081 GACTCGCAGAGGCTTTGCCATGA 1104
Db 1081 GACTCGCAGAGGCTTTGCCATGA 1104

RESULT 3

US-10-222-075-2
; Sequence 2, Application US/10222075
; Publication No. US20040033606A1
; GENERAL INFORMATION:
; APPLICANT: Zou, Jitao
; TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE
; FILE REFERENCE: 3015-5442US
; CURRENT APPLICATION NUMBER: US/10/222.075
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Brassica rapa
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: PDHK cDNA from B. rapa
US-10-222-075-2

Query Match 99.2%; Score 1095.6; DB 17; Length 1104;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1098; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 1 ATGGCGGTGAAGAGGCTTACGAGATGTTTTCGAAGCTTGTGATCGAGAGCTTACAGA 60
Db 1 ATGGCGGTGAAGAGGCTTACGAGATGTTTTCGAAGCTTGTGATCGAGAGCTTACAGA 60
Qy 61 TGGGAGATGATGAAGAGGCGGCTGAGCTCAGATCAATGATGAGTTCCGTTCCACT 120
Db 61 TGGGAGATGATGAAGAGGCGGCTGAGCTCAGATCAATGATGAGTTCCGTTCCACT 120
Qy 121 CCCACTGAGAGAACTTCTGATCTCGGCGCAGTTTCTTCAACAAGAGCTTCCGATTGG 180
Db 121 CCCACTGAGAGAACTTCTGATCTCGGCGCAGTTTCTTCAACAAGAGCTTCCGATTGG 180
Qy 181 ATGCGCAGGCGTGCATCGAATCGAGACCTGCTTATGCGCTCTCTGAGAAACCTGCC 240
Db 181 ATGCGCAGGCGTGCATCGAATCGAGACCTGCTTATGCGCTCTCTGAGAAACCTGCC 240
Qy 241 GTCTTGAAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 241 GTCTTGAAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Qy 301 ATCAAGATTAATCTGCTGATGAGAAAGAGTTGACACAGATGATCAAGGCTTTAAAGTAAG 360

Db 301 ATCAAGATTAATCTGCTGATGAGAAAGAGTTGACACAGATGATTAAGGCTGTTAAAGTAAG 360
Qy 361 CACAAACAAGTGGTCCCATGATGAGCTCTGGGTGTGAACCACTGAAAGAAAGTAAGAA 420
Db 361 CACAAACAAGTGGTCCCATGATGAGCTCTGGGTGTGAACCACTGAAAGAAAGTAAGAA 420
Qy 421 CTCTACGAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 421 CTCTACGAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Qy 481 GGGATCCGTAATGCTTATCGGAGAGATGTTGAGTTGATGATGATGATGATGATGATGAT 540
Db 481 GGGATCCGTAATGCTTATCGGAGAGATGTTGAGTTGATGATGATGATGATGATGATGAT 540
Qy 541 ACAGTGGTAACTATACACACCAAGATGTCCTCTATGAGAGTGGCAGAGAAATGCTAGTAA 600
Db 541 ACAGTGGTAACTATACACACCAAGATGTCCTCTATGAGAGTGGCAGAGAAATGCTAGTAA 600
Qy 601 GATCAAGGTCGATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 601 GATCAAGGTCGATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 660
Qy 661 GATCAAGGTCGATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 661 GATCAAGGTCGATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 720
Qy 721 GTCAGAACTCTCTCCGCTGCTGCCAAGAGCGGTTTGTAACTGATTAAGGTTGCACA 780
Db 721 GTCAGAACTCTCTCCGCTGCTGCCAAGAGCGGTTTGTAACTGATTAAGGTTGCACA 780
Qy 781 CCAATCCGTAATCATGTTGCTGATGGAATGGAAGATGTTTAAATTAAGTCTCAGATGAA 840
Db 781 CCAATCCGTAATCATGTTGCTGATGGAATGGAAGATGTTTAAATTAAGTCTCAGATGAA 840
Qy 841 GGTGAGGTAATACCGAAGAGGGTCTCCCTAAATATTCCTTAACCTTAACAGACTGCA 900
Db 841 GGTGAGGTAATACCGAAGAGGGTCTCCCTAAATATTCCTTAACCTTAACAGACTGCA 900
Qy 901 AGAAACCCACTTGAAGAAGATGTGACTTGGGAAACCGGTATGTTCCCTGACTATGAGCT 960
Db 901 AGAAACCCACTTGAAGAAGATGTGACTTGGGAAACCGGTATGTTCCCTGACTATGAGCT 960
Qy 961 GGTATGATGATGCTGCTGCTATGATGCTGATGCTGCTGCTATGCTGCTGCTGCTGCTG 1020
Db 961 GGTATGATGATGCTGCTGCTATGATGCTGATGCTGCTGCTATGCTGCTGCTGCTGCTG 1020
Qy 1021 CAGATCATATCCATGGAAGATACGGGACTGATGCTTACTTGCATTTGCTGCTTGA 1080
Db 1021 CAGATCATATCCATGGAAGATACGGGACTGATGCTTACTTGCATTTGCTGCTTGA 1080
Qy 1081 GACTCGCAGAGGCTTTGCCATGA 1104
Db 1081 GACTCGCAGAGGCTTTGCCATGA 1104

RESULT 4

US-10-642-531-2
; Sequence 2, Application US/10642531
; Publication No. US20040248302A1
; GENERAL INFORMATION:
; APPLICANT: Marillia, Elizabeth-France
; APPLICANT: Zou, Jitao
; TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE
; FILE REFERENCE: 3015-5442US
; CURRENT APPLICATION NUMBER: US/10/642.531
; PRIOR APPLICATION NUMBER: 2003-08-15
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2

; LENGTH: 1104
 ; TYPE: DNA
 ; ORGANISM: Brassica rapa
 ; FEATURE:
 ; OTHER INFORMATION: PDHK cDNA from B. rapa
 US-10-642-531-2

Query Match 99.2%; Score 1095.6; DB 18; Length 1104;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1098; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 ATGGCGGTGAAGAGGCTAGCAGATGTTTGAAGAGCTTGATCGAGAGCTTACAGA 60
DB 1 ATGGCGGTGAAGAGGCTAGCAGATGTTTGAAGAGCTTGATCGAGAGCTTACAGA 60
QY 61 TGGGGATGCATGACAGACGCGGCGTGAAGCTCAGTACATGATGAGATGGTTCCACT 120
DB 61 TGGGGATGCATGACAGACGCGGCGTGAAGCTCAGTACATGATGAGATGGTTCCACT 120
QY 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAAGAGACTTCCGATTCGG 180
DB 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAAGAGACTTCCGATTCGG 180
QY 181 ATCGGAGGCGTGCATGAACTCGAGACGCTGCTTATGCGCTCTCTGAGAAACCTGCC 240
DB 181 ATCGGAGGCGTGCATGAACTCGAGACGCTGCTTATGCGCTCTCTGAGAAACCTGCC 240
QY 241 GTCTTGAAGTAAAGATTTGGTATGTGAGTCAATTCAGGACATGAGAGCGTTCCGAG 300
DB 241 GTCTTGAAGTAAAGATTTGGTATGTGAGTCAATTCAGGACATGAGAGCGTTCCGAG 300
QY 301 ATCAAGATATCTGCTGATGAGAAAGATTCAACAGATGATCAAGGCTGTTAAAGTAAG 360
DB 301 ATCAAGATATCTGCTGATGAGAAAGATTCAACAGATGATCAAGGCTGTTAAAGTAAG 360
QY 361 CACAACAACGTGTGTTCCCATGATGAGGCTCTGGGTGAAACCAAGCTGAGAAAGAAATGAA 420
DB 361 CACAACAACGTGTGTTCCCATGATGAGGCTCTGGGTGAAACCAAGCTGAGAAAGAAATGAA 420
QY 421 CTCTACGAAAGCTTGAATGAGATTCATCAGTTTCTTGATGCGCTTCTGCTGCTGATA 480
DB 421 CTCTACGAAAGCTTGAATGAGATTCATCAGTTTCTTGATGCGCTTCTGCTGCTGATA 480
QY 481 GGGATCGGTATGCTTATCGGCGAGCATGTTGATGATGATGATGATGATGATGATGATG 540
DB 481 GGGATCGGTATGCTTATCGGCGAGCATGTTGATGATGATGATGATGATGATGATGATG 540
QY 541 ACAAGTGGTATACATACACACCAAGATGCTCTATGAGAGTGGCAAGAAATGCTAGTGA 600
DB 541 ACAAGTGGTATACATACACCAAGATGCTCTATGAGAGTGGCAAGAAATGCTAGTGA 600
QY 601 GATGCAAGTGCATGTTGTTTCAAGAGTATGTTGCTGCTCGGAGATTAACATATATGCG 660
DB 601 GATGCAAGTGCATGTTGTTTCAAGAGTATGTTGCTGCTCGGAGATTAACATATATGCG 660
QY 661 GATCACAAGTTCATTTCCGTATGTTCCGACCATTTGATGATGATGATGATGATGATG 720
DB 661 GATCACAAGTTCATTTCCGTATGTTCCGACCATTTGATGATGATGATGATGATGATG 720
QY 721 GTCAAGAACTCTCCGTGCTGTCACAGACGCGTTGTTGATGCTGATAGGTTGACCA 780
DB 721 GTCAAGAACTCTCCGTGCTGTCACAGACGCGTTGTTGATGCTGATAGGTTGACCA 780
QY 781 CCAATCGGTATCTGTTGCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 840
DB 781 CCAATCGGTATCTGTTGCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 840
QY 841 GGTGAGGTATACCGAGAGCGGTCTCCCTAAATATTCATTAAGCTTACAGCACTGCA 900
DB 841 GGTGAGGTATACCGAGAGCGGTCTCCCTAAATATTCATTAAGCTTACAGCACTGCA 900
QY 901 AGAAACCCACTGAGAAAGATGAGACTTGGGAACCGCTGATGCTCCCTGATAGGCT 960
DB 901 AGAAACCCACTGAGAAAGATGAGACTTGGGAACCGCTGATGCTCCCTGATAGGCT 960
  
```

```

DB 901 AGAAACCCACTGAGAAAGATGAGACTTGGGAACCGCTGATGCTCCCTGATAGGCT 960
QY 961 GGTATGTTATGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB 961 GGTATGTTATGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1020
QY 1021 CAGATCATATCATGAGAGATACGGGACTGATGCTTACTTGCATTTGCTGCTTGA 1080
DB 1021 CAGATCATATCCATGAGAGATACGGGACTGATGCTTACTTGCATTTGCTGCTTGA 1080
QY 1081 GACTCGAGAGGCTTGGCATGA 1104
DB 1081 GACTCGAGAGGCTTGGCATGA 1104
  
```

```

RESULT 5
US-10-222-075-4
; Sequence 4, Application US/10222075
; Publication No. US20040033606A1
; GENERAL INFORMATION:
; APPLICANT: Zou, Jitao
; TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE
; FILE REFERENCE: 3015-5442US
; CURRENT APPLICATION NUMBER: US/10/222,075
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Brassica carinata
; NAME/KEY: misc feature
; OTHER INFORMATION: PDHK cDNA from B. carinata
US-10-222-075-4
  
```

Query Match 96.2%; Score 1062.4; DB 17; Length 1104;
 Best Local Similarity 97.6%; Pred. No. 0;
 Matches 1078; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```

QY 1 ATGGCGGTGAAGAGGCTAGCAGATGTTTGAAGAGCTTGATCGAGAGCTTACAGA 60
DB 1 ATGGCGGTGAAGAGGCTAGCAGATGTTTGAAGAGCTTGATCGAGAGCTTACAGA 60
QY 61 TGGGGATGCATGACAGACGCGGCGTGAAGCTCAGTACATGATGAGATGGTTCCACT 120
DB 61 TGGGGATGCATGACAGACGCGGCGTGAAGCTCAGTACATGATGAGATGGTTCCACT 120
QY 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAAGAGACTTCCGATTCGG 180
DB 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAAGAGACTTCCGATTCGG 180
QY 181 ATCGGAGGCGTGCATGAACTCGAGACGCTGCTTATGCGCTCTCTGAGAAACCTGCC 240
DB 181 ATCGGAGGCGTGCATGAACTCGAGACGCTGCTTATGCGCTCTCTGAGAAACCTGCC 240
QY 241 GTCTTGAAGTAAAGATTTGGTATGTGAGTCAATTCAGGACATGAGAGCGTTCCGAG 300
DB 241 GTCTTGAAGTAAAGATTTGGTATGTGAGTCAATTCAGGACATGAGAGCGTTCCGAG 300
QY 301 ATCAAGATATCTGCTGATGAGAAAGATTCAACAGATGATCAAGGCTGTTAAAGTAAG 360
DB 301 ATCAAGATATCTGCTGATGAGAAAGATTCAACAGATGATCAAGGCTGTTAAAGTAAG 360
QY 361 CACAACAACGTGTGTTCCCATGATGAGGCTCTGGGTGAAACCAAGCTGAGAAAGAAATGAA 420
DB 361 CACAACAACGTGTGTTCCCATGATGAGGCTCTGGGTGAAACCAAGCTGAGAAAGAAATGAA 420
QY 421 CTCTACGAAAGCTTGAATGAGATTCATCAGTTTCTTGATGCGCTTCTGCTGCTGATA 480
DB 421 CTCTACGAAAGCTTGAATGAGATTCATCAGTTTCTTGATGCGCTTCTGCTGCTGATA 480
QY 481 GGGATCGGTATGCTTATCGGCGAGCATGTTGATGATGATGATGATGATGATGATGATG 540
DB 481 GGGATCGGTATGCTTATCGGCGAGCATGTTGATGATGATGATGATGATGATGATGATG 540
  
```

Db 481 GGGATCGGTATGCTTATCGGGCAGCATGTTGATGATTAATCCAAACCCACACTTAC 540
Qy 541 ACAGTGGTTACATACACACCAAGATGCTCTTATGAGAGTGGCAAGAAATGCTAGTGA 600
Db 541 ACTGTGGTTACATACACACCAAGATGCTCTTATGAGAGTGGCAAGAAATGCTAGTGA 600
Qy 601 GATGCAAGTGCATGTTGTTCCAGAGATGATGTTCTGCTCCGAGATAAACATATATGCG 660
Db 601 GATGCAAGTGCATGTTGTTCCAGAGATGATGTTCTGCTCCGAGATAAACATATATGCG 660
Qy 661 GATCCAAAGTTCACTTTCCGATGTTCCGACCCATTTGACCTTATGTTAGTGA 720
Db 661 GATCCAAAGTTCACTTTCCGATGTTCCGACCCATTTGACCTTATGTTAGTGA 720
Qy 721 GTCAGAACTCTCTCCGCTGCTGCTCCAGAGCGGTTTGTGATCTGATAGGTTGACCA 780
Db 721 GTCAGAACTCTCTACGCTGCTGCTCCAGAGCGGTTTGTGATCTGATAGGTTGACCA 780
Qy 781 CCAATCCGATTCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 781 CCAATCCGATTCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Qy 841 GGTGAGATATACCGAGAGCGGCTCTCCATAATTTCACTTACAGCATGCA 900
Db 841 GGTGAGATATACCGAGAGCGGCTCTCCATAATTTCACTTACAGCATGCA 900
Qy 901 AGAAACCCACTTGAAGAAAGATGAGACTTGGGAAACCGCTGATGTTCCCGTATAGCT 960
Db 901 AGAAACCCGCTTGAAGAAAGATGAGACTTGGGAAACCGCTGATGTTCCCGTATAGCT 960
Qy 961 GGTATAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 961 GGTATAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy 1021 CAGATCATATCCATGAGAGATACGGGATGATGCTTACTTGCATCTTGTCTGCTTGA 1080
Db 1021 CAGATCATATCCATGAGAGATACGGGATGATGCTTACTTGCATCTTGTCTGCTTGA 1080
Qy 1081 GACTCGAGAGAGCTTTGCCATGA 1104
Db 1081 GACTCGAGAGAGCTTTGCCATGA 1104
RESULT 6
US-10-642-531-4
Sequence 4, Application US/10642531
Publication No. US20040248302A1
GENERAL INFORMATION:
APPLICANT: Marillia, Elizabeth-France
APPLICANT: Zou, Jitao
APPLICANT: Taylor, David C.
TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE
FILE REFERENCE: 3015-5442US
CURRENT APPLICATION NUMBER: US/10/642,531
PRIOR FILING DATE: 2003-08-15
PRIOR APPLICATION NUMBER: 10/222,075
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 1104
TYPE: DNA
ORGANISM: Brassica carinata
FEATURE:
OTHER INFORMATION: PDHK cDNA from B. carinata
US-10-642-531-4
Query Match 96.2%; Score 1062.4; DB 18; Length 1104;
Best Local Similarity 97.6%; Pred. No. 0;
Matches 1078; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy 1 ATGCGCGTGAAGAGGCTAGCGAGATGTTTCAGAGAGCTTGATCGAGAGCTTCAAGA 60

Db 1 ATGCGCGTGAAGAGGCTAGCGAGATGTTTCAGAGAGCTTGATCGAGAGCTTCAAGA 60
Qy 61 TGGGATGATGAAGAGAGGCGCTGAGCTCAGAGTACATGATGAGATTGCTGCTCACT 120
Db 61 TGGGATGATGAAGAGAGGCGCTGAGCTCAGAGTACATGATGAGATTGCTGCTCACT 120
Qy 121 CCCACTGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAGACTTCCGATTCCG 180
Db 121 CCCACTGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAGACTTCCGATTCCG 180
Qy 181 ATGCGAGGCGTGCATGCAATCGAGACGCTCTTATGCTCTGAGAAACCTGCC 240
Db 181 ATGCGAGGCGTGCATGCAATCGAGACGCTCTTATGCTCTGAGAAACCTGCC 240
Qy 241 GTCCTGAAGTGAAGATGTTGTTATGAGATCTTACAGGACATGAGAGGTTTCTGAG 300
Db 241 GTCCTGAAGTGAAGATGTTGTTATGAGATCTTACAGGACATGAGAGGTTTCTGAG 300
Qy 301 ATCAAGATPACTGATGATGAGAAAGATTCAACAGATGATCAAGGCTGTTAAAGTAA 360
Db 301 ATCAAGATPACTGATGATGAGAAAGATTCAACAGATGATCAAGGCTGTTAAAGTAA 360
Qy 361 CACAAACGCTGCTCCATGATGCTCTGGGTGTAACCACTGAAGAAAGATGAA 420
Db 361 CACAAACGCTGCTCCATGATGCTCTGGGTGTAACCACTGAAGAAAGATGAA 420
Qy 421 CTCTAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 421 CTCTAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Qy 481 GGGATCCGATGCTTATCGGGCAGAGTGTGAGTTGATGATGATGATGATGATGATGATGAT 540
Db 481 GGGATCCGATGCTTATCGGGCAGAGTGTGAGTTGATGATGATGATGATGATGATGATGAT 540
Qy 541 ACAGTGGTTACATACACACCAAGATGCTCTTATGAGAGTGGCAAGAAATGCTAGTGA 600
Db 541 ACAGTGGTTACATACACACCAAGATGCTCTTATGAGAGTGGCAAGAAATGCTAGTGA 600
Qy 601 GATGCAAGTGCATGTTGTTCCAGAGATGATGTTCTGCTCCGAGATAAACATATATGCG 660
Db 601 GATGCAAGTGCATGTTGTTCCAGAGATGATGTTCTGCTCCGAGATAAACATATATGCG 660
Qy 661 GATCCAAAGTTCACTTTCCGATGTTCCGACCCATTTGACCTTATGTTAGTGA 720
Db 661 GATCCAAAGTTCACTTTCCGATGTTCCGACCCATTTGACCTTATGTTAGTGA 720
Qy 721 GTCAGAACTCTCTCCGCTGCTGCTCCAGAGCGGTTTGTGATCTGATAGGTTGACCA 780
Db 721 GTCAGAACTCTCTACGCTGCTGCTCCAGAGCGGTTTGTGATCTGATAGGTTGACCA 780
Qy 781 CCAATCCGATTCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 781 CCAATCCGATTCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Qy 841 GGTGAGATATACCGAGAGCGGCTCTCCATAATTTCACTTACAGCATGCA 900
Db 841 GGTGAGATATACCGAGAGCGGCTCTCCATAATTTCACTTACAGCATGCA 900
Qy 901 AGAAACCCACTTGAAGAAAGATGAGACTTGGGAAACCGCTGATGTTCCCGTATAGCT 960
Db 901 AGAAACCCGCTTGAAGAAAGATGAGACTTGGGAAACCGCTGATGTTCCCGTATAGCT 960
Qy 961 GGTATAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 961 GGTATAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy 1021 CAGATCATATCCATGAGAGATACGGGATGATGCTTACTTGCATCTTGTCTGCTTGA 1080
Db 1021 CAGATCATATCCATGAGAGATACGGGATGATGCTTACTTGCATCTTGTCTGCTTGA 1080
Qy 1081 GACTCGAGAGAGCTTTGCCATGA 1104
Db 1081 GACTCGAGAGAGCTTTGCCATGA 1104

Db 1081 GACTCGCAGAGCCTTGCATGA 1104

RESULT 7

US-10-222-075-3

Sequence 3, Application US/10222075

Publication No. US2004003606A1

GENERAL INFORMATION:

APPLICANT: Zou, Jitao

TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE

FILE REFERENCE: 3015-5442US

CURRENT APPLICATION NUMBER: US/10/222,075

CURRENT FILING DATE: 2002-08-16

NUMBER OF SEQ ID NOS: 15

SOFTWARE: Patentin version 3.1

SEQ ID NO 3

LENGTH: 1104

TYPE: DNA

ORGANISM: Brassica oleracea

FEATURE:

NAME/KEY: misc. feature

OTHER INFORMATION: PDHK cDNA from B. oleracea

US-10-222-075-3

Query Match 96.2%; Score 1062; DB 17; Length 1104;

Best Local Similarity 97.0%; Pred. No. 0;

Matches 1071; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGGCGGTGAAGAGGCTAGCGAGATGTTTCGAAGAGCCTTGATCGAGAGCTTCACAGA 60

Db 1 ATGGCGGTGAAGAGGCTAGCGAGATGTTTCGAAGAGCCTTGATCGAGAGCTTCACAGA 60

QY 61 TGGGGATGCATGAAGCAGACGGCGGTGAGCCTCAGGTACATGATGAGATTGCTCCACT 120

Db 61 TGGGGATGCATGAAGCAGACGGCGGTGAGCCTCAGGTACATGATGAGATTGCTCCACT 120

QY 121 CCCACTGAGAGAAACCTTCGATCTCGGCGCAGTTTCTTCAAGAGCCTTCGATTCGG 180

Db 121 CCCACTGAGAGAAACCTTCGATCTCGGCGCAGTTTCTTCAAGAGCCTTCGATTCGG 180

QY 121 CCCACTGAGAGAAACCTTCGATCTCGGCGCAGTTTCTTCAAGAGCCTTCGATTCGG 180

Db 121 CCCACTGAGAGAAACCTTCGATCTCGGCGCAGTTTCTTCAAGAGCCTTCGATTCGG 180

QY 181 ATCGCAGAGCGTGGCGATGAACTCGAGACGCTGCTTATGGCCTCTCTGAGAAACCTGCC 240

Db 181 ATCGCAGAGCGTGGCGATGAACTCGAGACGCTGCTTATGGCCTCTCTGAGAAACCTGCC 240

QY 241 GCTTGAAGGTAAAGATGGTATGTGAGTCAATTCAGGACATGAGAGCGTTTCTGAG 300

Db 241 GCTTGAAGGTAAAGATGGTATGTGAGTCAATTCAGGACATGAGAGCGTTTCTGAG 300

QY 241 GCTTGAAGGTAAAGATGGTATGTGAGTCAATTCAGGACATGAGAGCGTTTCTGAG 300

Db 241 GCTTGAAGGTAAAGATGGTATGTGAGTCAATTCAGGACATGAGAGCGTTTCTGAG 300

QY 301 ATCAAGGATATCTGCTGATGAGAAAGATTCACACAGATGATCAAGGCTGTTAAAGTAAG 360

Db 301 ATCAAGGATATCTGCTGATGAGAAAGATTCACACAGATGATCAAGGCTGTTAAAGTAAG 360

QY 361 CACAAACAAGTGTTCCTCATGATGGCTCTGGGTGTGAACCAAGTGAAGAAAGATGAA 420

Db 361 CACAAACAAGTGTTCCTCATGATGGCTCTGGGTGTGAACCAAGTGAAGAAAGATGAA 420

QY 361 CACAAACAAGTGTTCCTCATGATGGCTCTGGGTGTGAACCAAGTGAAGAAAGATGAA 420

Db 361 CACAAACAAGTGTTCCTCATGATGGCTCTGGGTGTGAACCAAGTGAAGAAAGATGAA 420

QY 421 CTCTACGAAAGCTGTGATGAGATTCAATCAGTTCTTGATGCTTCACTATGTTCTGATTA 480

Db 421 CTCTACGAAAGCTGTGATGAGATTCAATCAGTTCTTGATGCTTCACTATGTTCTGATTA 480

QY 421 CTCTACGAAAGCTGTGATGAGATTCAATCAGTTCTTGATGCTTCACTATGTTCTGATTA 480

Db 421 CTCTACGAAAGCTGTGATGAGATTCAATCAGTTCTTGATGCTTCACTATGTTCTGATTA 480

QY 481 GGGATCCGTATGCTTATCGGGCAGCATGTTGAGTTGCATATCAAAACCAACCACTTCA 540

Db 481 GGGATCCGTATGCTTATCGGGCAGCATGTTGAGTTGCATATCAAAACCAACCACTTCA 540

QY 481 GGGATCCGTATGCTTATCGGGCAGCATGTTGAGTTGCATATCAAAACCAACCACTTCA 540

Db 481 GGGATCCGTATGCTTATCGGGCAGCATGTTGAGTTGCATATCAAAACCAACCACTTCA 540

QY 541 ACAGTGGGTAACTATACACCAAGATGCTCTCTATGAGAGTGGCAAGATGCTATGTA 600

Db 541 ACAGTGGGTAACTATACACCAAGATGCTCTCTATGAGAGTGGCAAGATGCTATGTA 600

QY 541 ACAGTGGGTAACTATACACCAAGATGCTCTCTATGAGAGTGGCAAGATGCTATGTA 600

Db 541 ACAGTGGGTAACTATACACCAAGATGCTCTCTATGAGAGTGGCAAGATGCTATGTA 600

QY 601 GATGCAAGGTCAATTTGTTCAAGAGATAGTGTCTCGCGAGATTAACATATATGTC 660

Db 601 GATGCAAGGTCAATTTGTTCAAGAGATAGTGTCTCGCGAGATTAACATATATGTC 660

QY 601 GATGCAAGGTCAATTTGTTCAAGAGATAGTGTCTCGCGAGATTAACATATATGTC 660

Db 601 GATGCAAGGTCAATTTGTTCAAGAGATAGTGTCTCGCGAGATTAACATATATGTC 660

QY 661 GATCCAGTTTCACTTTCCGTATGTTCCGACCACTTTCATTTATGCTATGATGATTA 720

Db 661 GATCCAGTTTCACTTTCCGTATGTTCCGACCACTTTCATTTATGCTATGATGATTA 720

Db 661 GATCCAGTTTCACTTTCCGTATGTTCCGACCACTTTCATTTATGCTATGATGATTA 720

QY 721 GTCAAGAACTCTCTCGGTGCTGTCCAAAGCGGTTTGTGACCTCTATAGGTTGACCA 780

Db 721 GTCAAGAACTCTCTCGGTGCTGTCCAAAGCGGTTTGTGACCTCTATAGGTTGACCA 780

QY 781 CCAATCCGTATCAATTTGCTGATGAGATGCAATGATTAACAATAAGGTTCAAGTGA 840

Db 781 CCAATCCGTATCAATTTGCTGATGAGATGCAATGATTAACAATAAGGTTCAAGTGA 840

QY 781 CCAATCCGTATCAATTTGCTGATGAGATGCAATGATTAACAATAAGGTTCAAGTGA 840

Db 781 CCAATCCGTATCAATTTGCTGATGAGATGCAATGATTAACAATAAGGTTCAAGTGA 840

QY 841 GGTGAGGTATACCGAAGACCGCTCTCCCTAAATTTACTTACTTCAACGACTGCA 900

Db 841 GGTGAGGTATACCGAAGACCGCTCTCCCTAAATTTACTTACTTCAACGACTGCA 900

QY 841 GGTGAGGTATACCGAAGACCGCTCTCCCTAAATTTACTTACTTCAACGACTGCA 900

Db 841 GGTGAGGTATACCGAAGACCGCTCTCCCTAAATTTACTTACTTCAACGACTGCA 900

QY 901 AGAACCACCTGAGAGAGATGAGACTTGGGAAACCGCTGATGTTCCCTGACTATGGCT 960

Db 901 AGAACCACCTGAGAGAGATGAGACTTGGGAAACCGCTGATGTTCCCTGACTATGGCT 960

QY 901 AGAACCACCTGAGAGAGATGAGACTTGGGAAACCGCTGATGTTCCCTGACTATGGCT 960

Db 901 AGAACCACCTGAGAGAGATGAGACTTGGGAAACCGCTGATGTTCCCTGACTATGGCT 960

QY 961 GGTATGTTATGCTGCTCTATTTAGTGGCTTGTATGCTGCTATTTTGTGAGATTTG 1020

Db 961 GGTATGTTATGCTGCTCTATTTAGTGGCTTGTATGCTGCTATTTTGTGAGATTTG 1020

QY 961 GGTATGTTATGCTGCTCTATTTAGTGGCTTGTATGCTGCTATTTTGTGAGATTTG 1020

Db 961 GGTATGTTATGCTGCTCTATTTAGTGGCTTGTATGCTGCTATTTTGTGAGATTTG 1020

QY 1021 CAGATCATATCCATGAGAGATACGGAAGTATGCTTACTTGCATTTGCTCTTGTGA 1080

Db 1021 CAGATCATATCCATGAGAGATACGGAAGTATGCTTACTTGCATTTGCTCTTGTGA 1080

QY 1081 GACTCGCAGAGCCTTTGCCATGA 1104

Db 1081 GACTCGCAGAGCCTTTGCCATGA 1104

RESULT 8

US-10-642-531-3

Sequence 3, Application US/10642531

Publication No. US20040248302A1

GENERAL INFORMATION:

APPLICANT: Marillia, Elizabeth-France

APPLICANT: Zou, Jitao

TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE

FILE REFERENCE: 3015-5442US

CURRENT APPLICATION NUMBER: US/10/642,531

PRIOR APPLICATION NUMBER: 2003-08-15

PRIOR FILING DATE: 10/222,075

NUMBER OF SEQ ID NOS: 19

SOFTWARE: Patentin version 3.1

SEQ ID NO 3

LENGTH: 1104

TYPE: DNA

ORGANISM: Brassica oleracea

FEATURE:

OTHER INFORMATION: PDHK cDNA from B. oleracea

US-10-642-531-3

Query Match 96.2%; Score 1062; DB 18; Length 1104;

Best Local Similarity 97.0%; Pred. No. 0;

Matches 1071; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGGCGGTGAAGAGGCTAGCGAGATGTTTCGAAGAGCCTTGATCGAGAGCTTCACAGA 60

Db 1 ATGGCGGTGAAGAGGCTAGCGAGATGTTTCGAAGAGCCTTGATCGAGAGCTTCACAGA 60

QY 61 TGGGGATGCATGAAGCAGACGGCGGTGAGCCTCAGGTACATGATGAGATTGCTCCACT 120

Db 61 TGGGGATGCATGAAGCAGACGGCGGTGAGCCTCAGGTACATGATGAGATTGCTCCACT 120

QY 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAGAGCTTCGATTCGG 180

Db 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAGAGCTTCGATTCGG 180

QY 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAGAGCTTCGATTCGG 180

Db 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAGAGCTTCGATTCGG 180

QY 181 ATCGCAGAGCGTGGCGATGAACTCGAGACGCTGCTTATGGCCTCTCTGAGAAACCTGCC 240

Db 181 ATCGCAGAGCGTGGCGATGAACTCGAGACGCTGCTTATGGCCTCTCTGAGAAACCTGCC 240

```

Db 181 ATGCGAGGCGTGCGATCGAACTCGAGACCGCTGCTTATGCGCTCTCTGAGAAACCTGCC 240
Qy 241 GTCTTGAAGGTAGAGATTGATATGAGATCATTTGAGGACATGAGAGGTTTCTCGAG 300
Db 241 GTCTTGAAGGTAGAGATTGATATGAGATCATTTGAGGACATGAGAGGTTTCTCGAG 300
Qy 301 ATCAAGATATCTCTGATGAGAAAGAGTTCACACAGATGATCAAGGCTGTTAAAGTAGG 360
Db 301 ATCAAGATATCTCTGATGAGAAAGAGTTCACACAGATGATCAAGGCTGTTAAAGTAGG 360
Qy 361 CACACAAACGCTGTTCCATGATGCTCTGAGTGTGAACCACTGAAGAAAGAAATGAAA 420
Db 361 CACACAAACGCTGTTCCATGATGCTCTGAGTGTGAACCACTGAAGAAAGAAATGAAA 420
Qy 421 CTCTACGAAAGAGTGTGATGATTCATCACTTTCTTATATGCTTCTTCTCTCGATA 480
Db 421 CTCTACGAAAGAGTGTGATGATTCATCACTTTCTTATATGCTTCTTCTCTCGATA 480
Qy 481 GGGATCCGATGCTTATCGGGCAGCATGTTGAGTGCATATCCAAACCCACACTTCAC 540
Db 481 GGGATCCGATGCTTATCGGGCAGCATGTTGAGTGCATATCCAAACCCACACTTCAC 540
Qy 541 ACAGTGGTTACATACACACCAAGATGCTCTATGAGAGTGCAGAAATGCTATGAA 600
Db 541 ACTGGGTTACATACACCAAGATGCTCTATGAGAGTGCAGAAATGCTATGAA 600
Qy 601 GATGCAAGTGCATTTGTTTCAAGAGATATGCTTCTCTCGAGATTAACATATATGCG 660
Db 601 GATGCAAGTGCATTTGTTTCAAGAGATATGCTTCTCTCGAGATTAACATATATGCG 660
Qy 661 GATCCAAAGTTTCACTTTTCCGATATGCTCCAGACCATTTGATCTTATGATGATTA 720
Db 661 GATCCAAAGTTTCACTTTTCCGATATGCTCCAGACCATTTGATCTTATGATGATTA 720
Qy 721 GTCAAGAACTCTCTCGTGTGCTGCAAGACCGGTTTGTGATCTGATGAGGTTGACCA 780
Db 721 GTCAAGAACTCTCTCGTGTGCTGCAAGACCGGTTTGTGATCTGATGAGGTTGACCA 780
Qy 781 CCAATCCGATATCTTGTGCTGATGAAATGGAAGATTTCAATTAAGGCTCTAGATGAA 840
Db 781 CCAATCCGATATCTTGTGCTGATGAAATGGAAGATTTCAATTAAGGCTCTAGATGAA 840
Qy 841 GGTGAGGTATACCGAAGACGGCTCTCTAAATATTTCACTTACCTTACAGACTGCA 900
Db 841 GGTGAGGTATACCGAAGACGGCTCTCTAAATATTTCACTTACCTTACAGACTGCA 900
Qy 901 AGAAACCCCTTGAAGAGATGAGACTTGGGAAACCGCTGATGTTCCCTGACTATGCT 960
Db 901 AGAAACCCCTTGAAGAGATGAGACTTGGGAAACCGCTGATGTTCCCTGACTATGCT 960
Qy 961 GGTATAGTATATGCTGCTATTAATGCTGCTTGTATGCTGCTATTTGGTGAAGATTG 1020
Db 961 GGTATAGTATATGCTGCTATTAATGCTGCTTGTATGCTGCTATTTGGTGAAGATTG 1020
Qy 1021 CAGATCATATCATGAGAGATACGGGATGATGCTTACTTGAACCTTCTCGCTTGA 1080
Db 1021 CAGATCATATCATGAGAGATACGGGATGATGCTTACTTGAACCTTCTCGCTTGA 1080
Qy 1081 GACTGCGAAGAGCTTTTGGCATGA 1104
Db 1081 GACTGCGAAGAGCTTTTGGCATGA 1104

```

```

RESULT 9
US-10-202-428-1
; Sequence 1, Application US/10202428
; Publication No. US20030084472A1
; GENERAL INFORMATION:
; APPLICANT: Zou, Jitao
; APPLICANT: Taylor, David
; TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE
; FILE REFERENCE: 40942-1

```

```

; CURRENT APPLICATION NUMBER: US/10/202,428
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/09/355,912
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/CA98/00096
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/038,815
; PRIOR FILING DATE: 1997-02-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-202-428-1

Query Match 79.1%; Score 873.6; DB 14; Length 1457;
Best Local Similarity 87.6%; Pred. No. 2,9e-273;
Matches 967; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

Qy 1 ATGCGGTTGAAGAGCTAGCGAGATGTTTGGAGAGCTGATCGAGAGCTTCAAGA 60
Db 104 ATGCGAGTGAAGAGCTGCGAAATGTTCCGAAAGTTGATCGAAGATGTTCAAAA 163
Qy 61 TGGGATGATGAAGAGAGACGGCGTGAAGCTCAGGTACATGATGAGATTCAGT 120
Db 164 TGGGTTGATGAAGAGAAACCGGTGTTAGCTTATGATGATGAGATGAGTTGTTCCAA 223
Qy 121 CCACTGAGAGAAACCTTCTGATCTCGGCGAAGTTCTTCAAGAGCTTCCGATTGG 180
Db 224 CCACTGAGAGAAATCTTTGATTTCTGCTCAGCTTTTGTGATGAGAGCTTCCGATTGG 283
Qy 181 ATGCGAGGCGTGCATGCAATCGAGACGCTGCTATGAGGCTCTGAGAAACCTGCC 240
Db 284 GTGCGAGAGAGCATTCGAATCCAGACCTTCTTATGCTCTCTGATTAACCTGCC 343
Qy 241 GTCTTGAAGGTAGAGATTGATGAGATCATTTGAGAGTGCAGAGATGAGAGGTTTCTGAG 300
Db 344 GTTTGAAGGTGCGGATGTTGATTTGAAATCTTTCAAGGAGATGAGAGATTTCTGAG 403
Qy 301 ATCAAGATATCTGCTGATGAGAAAGTTCACACAGATGATCAAGGCTGTTAAAGTAGG 360
Db 404 ATTAAGATTCGGGTGATCGAGAGAAATTTCACTCAAGATGATTAAGGCTGCAAGTAGG 463
Qy 361 CACACAAACGCTGTTCCATGATGCTGCTGGGTGAGAACAGCTGAGAAAGAAATGAAA 420
Db 464 CATTAAGATGATGCTCCATGATGCTGCTGGGTGATTAAGCTTCAAGAAAGAAATGAAA 522
Qy 421 CTCTACGAAAGAGCTTATGATGATTCATCACTTTCTTATGCTTCTGCTATG 480
Db 523 --TTTGGAAATCTTATGATGATTCATCACTTTCTTATGCTTCTGCTGCAATC 580
Qy 481 GGGATCCGATATGCTTATCGGCGAGCATGTTGAGTGTGATTAATCCAAACCCACACTTCAC 540
Db 581 GGGATCCGATATGCTTATGAGGCGAGCATGTTGAGTGTGATTAATCCAAATCCACCGCTTCAT 640
Qy 541 ACAGTGGTTACATACACACCAAGATGCTCTTATGAGAGTGCAGAAATGCTATGAA 600
Db 641 ACAGTGGTTATATACACCAAGATGCTCTTATGAGAGTGCAGAAATGCTATGAA 700
Qy 601 GATGCAAGTGCATTTGTTTCAAGAGATGTTGCTTCTCGAGATTAACATATATGCG 660
Db 701 GATGCTCGGTCATTTGTTTCCAGAGTACGGTTTGCACCCGAAATTAACATATATGCG 760
Qy 661 GATCCAAAGTTTCACTTTTCCGATATGTTTCCGACCATTTGATCTTATGATGATTA 720
Db 761 GATCCCAAGTTTCACTTTTCCGATATGTTTCCGACCATTTGATCTTATGATGATGAGCTA 820
Qy 721 GTCAAGAACTCTCTCGGCTGCTCAAGAGCGGTTTGTGATCTGATAGAGTTGACCA 780
Db 821 GTCAAGAACTCTTACGCTGCTGCAAGAGCATTTGTTGATCTGATAGAGTTGACCA 880
Qy 781 CCAATCCGATATCTTGTGCTGATGAGATCGAAGATGTTCAATTAAGGCTTCAGATGAA 840

```


Db 1048 ACAATGGCTGATATGATATGCTCTTATTAGTCTCTATATAGCTCGATATTTTGA 1107
 Qy 1012 GGAGATTTGGAGATCATATCATATGAGATAGGAGCTGATGTTACTTGCATTTGCT 1071
 Db 1108 GGGATCTTCAAAATTAATCTTATGAGAAATGAGCACTGATGATATCTTCATTTGTCT 1167
 Qy 1072 CGTCTTGAGACTCGCAGAGAGCCTTGGCATGA 1104
 Db 1168 CGTTGGAGATTCACAGAAACCTTGGCTTGA 1200
 RESULT 11
 US-10-424-599-109588
 ; Sequence 109588, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yinhua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223) B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 109588
 ; LENGTH: 2028
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_69973C.1
 US-10-424-599-109588
 Query Match 50.7%; Score 559.4; DB 17; Length 2028;
 Best Local Similarity 70.9%; Pred. No. 9,7e-171;
 Matches 789; Conservative 0; Mismatches 311; Indels 13; Gaps 3;
 Qy 2 TGCGCGTGAAGAGGCTAGCGAGATGTTTTCGAAGAGCTTGATCGAGAGCTTCACAGAT 61
 Db 409 TGTGCACTACCAAGCACTGAGATTAATTTTCAAGTCTTCATATAGCTTGTCCAGAA 468
 Qy 62 GGGGATGATGAGAGAGAGAGCGGCGTGAAGCTCAGTACATATGATGAGT-TCGGTTCACT 120
 Db 469 CGGGTGTCTGAAGAGAGATGAGGATGATGTTGAAGGTCCTTGATGATGATGATGATGAT 528
 Qy 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAGCTTCCGATTCG 180
 Db 529 CCCATTGATCAAAAATGTGATGATGATGAGAGAGTTCACTGAAAATGATGAGGCGCATCA 588
 Qy 181 ATCGGAGGCGTGCATGCAACTCGAGAGCGCTGCTTATGGCCTCTGAGAAACCTGCC 240
 Db 589 AAGTGAAGCAGGGCCATGAGCTTGAGATCTTCCCTATGTTGTCTCAAAAACCTGCT 648
 Qy 241 GTCTTGAAGTAAAGATTTGATGTGAGTCAATTCAGGAGCATGAGAGCTTCTGAG 300
 Db 649 GTTTTAAAGTTAGGATTTGATGTGATGATCTTTCGATGATCTCAGAGCTTCCCCAAC 708
 Qy 301 ATCAAGGATACCTGCTGATGAGAAAGCTTCAACAAGATGATCAAGGCTGTTAAAGTAA 360
 Db 709 ATCAAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 768
 Qy 361 CACAACAGAGTGTCCCATGATGAGGCTGAGGAGTGAACAGAGTGAAGAAAGATGAA 420
 Db 769 CACAACAGAGTGTACCAACATGAGCTTGGGTTGTCAGCAATGAGAAAGATGATGAT 828
 Qy 421 C-----TCTACGAAAAGCTTGAAGATTCATCAGTTTCTTGAATGCTTCTAATTG 471
 Db 829 CCAAGATTTGTTATGAGATCTTGTGAGATTCATCAGATTTCTGAGCGGCTTCTAAG 888
 Qy 472 TCTCGTATAGGATCCGTATGCTTATTCGGGAGAGATTTGATGATGATGATGATGATGATGAT 531
 Db 889 TCAGGATTTGATCGGTATGCTTATTCGGGAGAGATTTGATGATGATGATGATGATGATGAT 948

Qy 532 CCACTTCAACAGAGTGTATACATACACCAAGATGCTCTATGAGAGTGGCAAGAAAT 591
 Db 949 TCTCCCACTGTTGTGGCTATATACATACAAAATGCTCTGTGAGAGTACAGAAAT 1008
 Qy 592 GCTAGTGAAGATGCAAGTGTGATTTTTCAGAGATGATGTTCTGCTCGGAGATTAAC 651
 Db 1009 GCCAGTGAAGATGCAAGTGTGATTTGTCGGAATATGAGATGCTGCTGATGAT 1068
 Qy 652 ATATATGAGATGCAAGTGTGATTTTTCAGAGATGATGTTCTGCTCGGAGATTAAC 711
 Db 1069 ATTTATGAGATGCAAGTGTGATTTTTCAGAGATGATGTTCTGCTCGGAGATTAAC 1128
 Qy 712 TATGATGATGCAAGTGTGATTTTTCAGAGATGATGTTCTGCTCGGAGATTAAC 771
 Db 1129 TTTGATGATGCAAGTGTGATTTTTCAGAGATGATGTTCTGCTCGGAGATTAAC 1188
 Qy 772 GTTGACACCAACATCCGTATGATTTTTCAGAGATGATGTTCTGCTCGGAGATTAAC 831
 Db 1189 GTTGACACCTCCATGATTAATGATTTTTCAGAGATGATGTTCTGCTCGGAGATTAAC 1248
 Qy 832 TCAGATGAGTGTGATGATTTTTCAGAGATGATGTTCTGCTCGGAGATTAAC 891
 Db 1249 TCAGATGAGTGTGATGATTTTTCAGAGATGATGTTCTGCTCGGAGATTAAC 1308
 Qy 892 AGCACTGCAAGAAACCACTTGAAGAGATGATGATTTTTCAGAGATGATGTTCTGCTCGGAGATTAAC 951
 Db 1309 AGTACTGCAAGAAACCACTTGAAGAGATGATGATTTTTCAGAGATGATGTTCTGCTCGGAGATTAAC 1365
 Qy 952 ACTAGTGTGATGATGATTTTTCAGAGATGATGTTCTGCTCGGAGATTAAC 1011
 Db 1366 ACATGCTGTGATGATGATTTTTCAGAGATGATGTTCTGCTCGGAGATTAAC 1425
 Qy 1012 GGAGATTTGCAAGTGTGATTTTTCAGAGATGATGTTCTGCTCGGAGATTAAC 1071
 Db 1426 GGGATCTTCAAAATTAATCTTATGAGAAATGAGCACTGATGATGATGATGATGATGAT 1485
 Qy 1072 CGTCTTGAGACTCGCAGAGAGCCTTGGCATGA 1104
 Db 1486 CGTTGGAGATTCACAGAAACCTTGGCTTGA 1200
 RESULT 12
 US-10-062-254-9
 ; Sequence 9, Application US/10062254
 ; Publication No. US20020138882A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cahoon, Edgar B
 ; APPLICANT: Cahoon, Rebecca E
 ; APPLICANT: Falco, Saverio Carl
 ; APPLICANT: Fang, Yiwen
 ; APPLICANT: Hanke, Sabine S.
 ; APPLICANT: Lee, Jian-Ming
 ; APPLICANT: Li, Zhongsen
 ; APPLICANT: Miao, Guo-Hua
 ; APPLICANT: Morgante, Michele
 ; APPLICANT: Niu, Xiping
 ; APPLICANT: Odell, Joan
 ; APPLICANT: Rafalski, Antoni
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Zheng, Peizhong
 ; APPLICANT: Zhu, Qun
 ; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/10/062,254
 ; PRIOR FILING DATE: 2002-02-01
 ; PRIOR APPLICATION NUMBER: 09/630,346
 ; PRIOR FILING DATE: 2000-07-28
 ; PRIOR APPLICATION NUMBER: 60/146511
 ; PRIOR FILING DATE: 1999-07-30
 ; PRIOR APPLICATION NUMBER: 60/156006
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: 60/156899

/ PRIOR FILING DATE: 1999-09-30
 / PRIOR APPLICATION NUMBER: 60/157287
 / PRIOR FILING DATE: 1999-10-01
 / PRIOR APPLICATION NUMBER: 60/169767
 / PRIOR FILING DATE: 1999-12-09
 / PRIOR APPLICATION NUMBER: 60/171054
 / PRIOR FILING DATE: 1999-12-16
 / PRIOR APPLICATION NUMBER: 60/172958
 / PRIOR FILING DATE: 1999-12-21
 / PRIOR APPLICATION NUMBER: 60/175151
 / PRIOR FILING DATE: 1999-12-22
 / PRIOR APPLICATION NUMBER: 60/173535
 / PRIOR FILING DATE: 1999-12-29
 / NUMBER OF SEQ ID NOS: 375
 / SOFTWARE: Microsoft Office 97
 / SEQ ID NO 9
 / LENGTH: 1481
 / TYPE: DNA
 / ORGANISM: *Oryza sativa*
 / US-10-062-254-9

Query Match 49.6%; Score 547.4; DB 13; Length 1481;
 Best Local Similarity 69.9%; Pred. No. 6.6e-167;
 Matches 778; Conservative 0; Mismatches 311; Indels 24; Gaps 2;

QY 4 GCGGTGAAGAAGCTTACGAGATGTTTCGAAAGCTTATCGAGACGTTCAACAGATGG 63
 DB 46 GGGGGGAGATGCGCGTCGAGCCGCGTGGCGAGGGCGGTCGAGAGAGTGGGAGATGG 105
 QY 64 GGATGCAATGAAGCAGACGCGGCTGAGCTTCAAGTACATGATGAGTTCGCTTCACTCC 123
 DB 106 GGGAGCATGAAGACGAGCGGGGTAGCGCTGCTACATGATGAGTTCGCGTCCGCC 165
 QY 124 ACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAGACTTCCGATTCCGATC 183
 DB 166 ACGGAGCGGAACCTGCTGCTCTCCGCCAGTTCCTGAGAGAGAGCTCCCAATCCGATC 225
 QY 184 GCGAGGCTGCGATCGAACTCGAGACGCTGCTTATGAGCTTCTGAGAAACCTGCGCTC 243
 DB 226 GCGGCGCGCGCTCGAGCTCGATCCCTCCCTTCCGCTCCCGCAAGCCCGCATC 285
 QY 244 TTGAAGGTAAAGATGATGATGAGTTCATTCAGGAGCATGAGAGCTTCTCGAGATC 303
 DB 286 CTGAGGTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 345
 QY 304 AAGGATATCTGCTGATGAGAAAGTTTCAACAGATGATGAGCTGTTAAAGTAAGCAGC 363
 DB 346 AAGGATATGAGAAAGCAGCTGCTTTCACGAGATGATGAGATGATGATGATGATGATG 405
 QY 364 AACAGCTGCTCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 412
 DB 406 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 465
 QY 413 -GAATGAACTTCAAGAAAGCTTGAATGAGATTCATCAAGTTTCTTATGCTCTTATCTG 471
 DB 466 ACAAGAGAAATCCCACTGATTTGAGAGATTCACAGAGTTTCTTATGAGATTTCTATG 525
 QY 472 TCTCGATAGGATCCGATGCTTATGCGGACAGATGTTGAGTTCATATCCAAACCA 531
 DB 526 TCAAGATTTGGATCCGATGCTTATGAGGACAGATGCTTATGATGATGATGATGATGAT 585
 QY 532 CCACTTCAACAGTGGTTACATACACCAAGATGCTTCATGAGAGTTCGCAAGAT 591
 DB 586 GAGCTGGCTCATAGGCTCATTAATACAGATTTCCCTATACAGATGCTCAAGCT 645
 QY 592 GCTAGTGAAGATCAAGGCTGATTTTCAAGAGATGATGCTTCTCGAGATTAAC 651
 DB 646 GCGAGTGAAGATCCGCTTATTTGTTGAGGAGATGATGATGATGATGATGATGATGATG 705
 QY 652 AATATAGGATCAAGTTCCTTCTGATGCTTCAAGATGATGATGATGATGATGATGATG 711
 DB 706 AATATAGGATCAAGTTCCTTCTGATGCTTCAAGATGATGATGATGATGATGATGATG 765

QY 712 TATGATTAAGTCAAGAACTCTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 771
 DB 766 TTGGAATGATGAAGAACTCTCGCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 825
 QY 772 GTTGACCCCAACCATCCGATATGATGCTGATGATGATGATGATGATGATGATGATGAT 831
 DB 826 GATGTTCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 885
 QY 832 TCAGATGAAGTGAAGTATACCGAAGACGCTTCCCTAAATATTAATTAATTAATTAAT 891
 DB 886 AGTATGAAGTGAAGTATACCGAAGACGCTTCCCTAAATATTAATTAATTAATTAAT 945
 QY 892 AGCATCTGAAGAAACCACTTGAAGAGATGAGTGGAAATGAGTGGAAATGAGTGGAAAT 951
 DB 946 AGCATCTGAAGAAATCCACT-----GATATGATGATGATGATGATGATGATGATG 993
 QY 952 ACTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1011
 DB 994 ACTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1053
 QY 1012 GAGATTTGCAATCATATCATATGAGAGATACGGAATGATGATGATGATGATGATGAT 1071
 DB 1054 GGTGACCTGCAATCATATCATATGAGAGATACGGAATGATGATGATGATGATGATGAT 1113
 QY 1072 CGTCTGAGACTCGAGAGAGCTTGGCATGA 1104
 DB 1114 CGTCTGAGACTCGAGAGAGCTTGGCATGA 1146

RESULT 13
 US-10-437-963-74020
 / Sequence 74020, Application US/10437963
 / Publication No. US2004012343A1
 / GENERAL INFORMATION:
 / APPLICANT: La Rosa, Thomas J.
 / APPLICANT: Kovalic, David K.
 / APPLICANT: Zhou, Yihua
 / APPLICANT: Cao, Yongwei
 / APPLICANT: Wu, Wei
 / APPLICANT: Boukhari, Andrew A.
 / APPLICANT: Barbazuk, Brad
 / APPLICANT: Li, Ping
 / TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 / FILE REFERENCE: 38-21(53221)B
 / CURRENT APPLICATION NUMBER: US/10/437,963
 / CURRENT FILING DATE: 2003-05-14
 / NUMBER OF SEQ ID NOS: 204966
 / SEQ ID NO 74020
 / LENGTH: 2892
 / TYPE: DNA
 / ORGANISM: *Oryza sativa*
 / FEATURE:
 / OTHER INFORMATION: Clone ID: PAT_MRT4530_7424C.1
 / US-10-437-963-74020

Query Match 49.6%; Score 547.4; DB 18; Length 2892;
 Best Local Similarity 69.9%; Pred. No. 9.8e-167;
 Matches 778; Conservative 0; Mismatches 311; Indels 24; Gaps 2;

QY 4 GCGGTGAAGAAGCTTACGAGATGTTTCGAAAGCTTATCGAGACGTTCAACAGATGG 63
 DB 1139 GGGGGGAGATGCGCGTCGAGCCGCGTGGCGAGGGCGGTCGAGAGAGTGGGAGATGG 1198
 QY 64 GGATGCAATGAAGCAGACGCGGCTGAGCTTCAAGTACATGATGAGTTCGCTTCACTCC 123
 DB 1199 GGGAGCATGAAGACGAGCGGGGTAGCGCTGCTACATGATGAGTTCGCGTCCGCC 1258
 QY 124 ACTGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAGAGCTTCGATTCGATC 183
 DB 1259 ACGGAGCGGAACCTGCTGCTCTCCGCGCATGTTCTCGAGAGAGAGCTCCCAATCCGATC 1318
 QY 184 GCGAGGCTGCGATCGAACTCGAGAGCGTGCCTTATGAGCTTCTTGAAGAAACCTGCGCTC 243


```

Db      1319 GCGGCGCGCGCTCGAGCTCGAGTCCCTCCCTCGGCGCTCCCGCAAGCCCGCATC 1378
Qy      244 TTGAAGTAAAGATTGGTATGTGAGTCAATTCAGGACATGAGAGCTTCTCGAATC 303
Db      1379 CTAAAGTGGGAGTGGTACTTGGAATCTTTCGCGCAATCAATCTTTCGCAAGTG 1438
Qy      304 AAGGATCTGCTGATGAGAAAGAGTTCACACAGATGATCAAGGCTGTTAAAGTAAAG 363
Db      1439 AGGAATAGGAACGACGAGCTCGCTTTTACGAGATGATCAAGATCAAGTACGCG 1498
Qy      364 AACACGTGTTCCCATGATGCTCTGGTGTGAACAGCTGAAGAAAG----- 412
Db      1499 AATTAATGTGTTCCAAAGATGCGACTGGAGTTCAACAGCTCAAGAACGAGAGTATG 1558
Qy      413 -GAATGAAACTCAACGAAAGCTTGAAGATTCATCAATTTCTTGATCGCTTACTTG 471
Db      1559 ACAAGAAAGATCCCGCTGATGAGAGATTCACAGATTTCTTGACAGATTTCTCAT 1618
Qy      472 TCTGTATAGGATCCGTATGCTTATCGGCGAGCATGTTGATGATTAATCCAAACCA 531
Db      1619 TCAGGATTTGGATCCGATGCTTATAGGCGAGATGCTTTCATGATCTGACCCA 1678
Qy      532 CCACTTCAACAGTGGTTCATACACACCAAGATGCTCTATGAGGTGGCAAGAT 591
Db      1679 GAGCTGGCGTCAATAGGCTCATTAATACAGAAATGTCCTATACAGTGGCTCAAG 1738
Qy      592 GCTAGTGAAGATCGAAGTGGATTTGTTTCAAGAGATGATGTTGCTCCGAGATTAAC 651
Db      1739 GCCAGTGAAGATCCCGCTTATTTGTTGAGGAAATGAGTCAAGCTCTGAGATTAAC 1798
Qy      652 ATATATGCGCATCAAGTTCATCTTTCGATGTTCCGACCATTTGATCTTATGATG 711
Db      1799 ATCTATGAGAACCCCACTTTTCAATTTCCATATGTTTCAATCAATCTACATCT 1858
Qy      712 TATGAGTATGTCAGAACTCTCTCCGCTGCTCCAGAGCGGTTGTTGATCTGATAG 771
Db      1859 TTTGAATGGTGAAGAACTCTTGGCGTACAGTACAGAAAGATATGAAATCCGATA 1918
Qy      772 GTTGACACCAACCAATCCGTATCATTTGTTGCTGATGGAATCGAAGTGTAAATPAA 831
Db      1919 GATGTTCTCTCAATTTGAATTTATGTTGATGAGGCAAGATGTAACATCAAGAT 1978
Qy      832 TCAGATGAAGGTGAAGGTATACGAGAAAGGCTCTCCCTAAATATTCATCACTTAC 891
Db      1979 AGTGAATGAAGTGTGGAATACCAAGAGCGGCTTCCAAAGATTTTCAATATCTAT 2038
Qy      892 AGCAGTGCAGAAACCACTTGAAGAAAGATGTGACTTGGAAACCGCTGATGTTCC 951
Db      2039 AGCAGTGCAGAAAGATCACT-----GATATGATGCTCCTAGTGAAGAGTA 2086
Qy      952 ACTATGAGCTGTTATGTTATGCTGCTGCTTATTAAGTGTGTTATGCTGCTATTT 1011
Db      2087 ACTATGAGCTGTTATGTTATGCTGCTTCAATCAATGCTGTTATGCTGATTTT 2146
Qy      1012 GGAGATTTGAGATTCATTCATGATGAGAGATTAACGAGCTGATGCTTACTTCACTT 1071
Db      2147 GGAGACTGCAATTCATCTCTATGAGAGATTCAGGACCGATGCTTACCTGCACTT 2206
Qy      1072 GGTCTTGAGACTCGCAGAGAGCTTTGCGATGA 1104
Db      2207 CGGTGGAGATTCGAGAGAGAGCTTTCCTTGA 2239

```

RESULT 14
US-10-425-114-15131

; Sequence 15131, Application US/10425114
; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jindong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

```

; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 15131
; LENGTH: 1165
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3051-054-H6_FLI
US-10-425-114-15131

```

Query Match

Best Local Similarity 75.5%; Pred. No. 1,9e-163;
Matches 697; Conservative 0; Mismatches 214; Indels 12; Gaps 2;

```

Qy      191 GTGCGATCGAAGCTCGAGAGCTGCTTATGCGCTCTGAGAAACCTGCGCTTGAAG 250
Db      2 GGGCATTTAGCTTGAAGATCTTCCCTATGCTTGTCTCAAAACCTGCTGTTTAAAG 61
Qy      251 TAAAGATTTGATATGATGATGATTCATTCAGGACATGAGAGCTTCTGATCAAGATA 310
Db      62 TTAGGATTTGATATGATGATTTCTTCGATGATCGAGAGCTTCCCAACATCAAGATG 121
Qy      311 CTGCTGATGAGAAAGATTCACACAGATGATCAAGCTGTTAAAGTAAAGCAACAACG 370
Db      122 TGAATGATGAAAGATTTCACTGAATGATCAAGGCAATCAAGTACAGGACACAATG 181
Qy      371 TGGTTCCTCATGATGCTGCTGGGTGTGAACAGCTGAAAGAAAGTAAAC----- 421
Db      182 TGGATCCCAAGAGGCTTGGGTGTGAACAGCTGAAAGTAAAGTAAAGTAAAGTAAAG 241
Qy      422 TCTACGAAAGCTTGAATGATGATTCATCATGTTTCTGATGCTTCTTACTTGTCTG 481
Db      242 TTTATGAAAGTCTTGTGATTCATCATGATTCATGATTCATGATTCATGATTCATG 301
Qy      482 GATTCGATGCTTATGCTGAGAGATGTTGATGATTCATTAATCCAAACCACTTACA 541
Db      302 GAATCCGATGCTTATGCTGAGAGATGTTGATGATTCATTAATCCAAACCACTTACA 361
Qy      542 CAGTGGTTCATCAACACCAAGATGCTCTATGAGAGGTGGAAGAGATGCTAGTAG 601
Db      362 TTTGGGCTATATATCAACAAAGATGCTCTGATGAGAGTGAAGAGAGTGAAG 421
Qy      602 ATGCAAGTCTGATTTGTTTCAAGAGATGATGCTGCTCCGAGATTAACATATATG 661
Db      422 ATGCAAGTCTGATTTGTTTCAAGAGATGATGCTGCTCCGAGATTAACATATATG 481
Qy      662 ATGCAAGTCTGATTTGTTTCAAGAGATGATGCTGCTCCGAGATTAACATATATG 721
Db      482 ATCTGATTTTACTTTTCCGATGTTTCAAGATGATGCTGCTCCGAGATTAACATAT 541
Qy      722 TCAAGAACTCTCTCCGCTGCTGCAAGAGGCTTGTGATGCTGATGAGGTTGCAAC 781
Db      542 TTAAGAACTCTCTCCGCTGCTGCAAGAGGCTTGTGATGCTGATGAGGTTGCAAC 601
Qy      782 CAATCGTATCATTTGTTGCTGATGATGCAAGATGTTCAATTAAGTCTCAGATAG 841
Db      602 CAATTAAGATTAAGTTGCTGATGATGCAAGATGTTCAATTAAGTCTCAGATAG 661
Qy      842 GTGAGGTATACGAGAGAGGCTCTCTTAAATATTCATTAACCTTACAGCATGCA 901
Db      662 GAGTGAATTTGCAAGAGAGGCTCTCTTAAATATTCATTAACCTTACAGCATGCA 721
Qy      902 GAAACCACTTGAAGAGATGATGCTGGAACCGCTGATGTTCCCTCACTAGTGGCT 961
Db      722 GAAACCACTTGAAGAGATGATGCTGGAACCGCTGATGTTCCCTCACTAGTGGCT 778
Qy      962 GTTATGTTATGATGCTGCTTATGATGCTGCTTATGATGCTGCTTATGATGAT 1021

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2005, 04:37:23 ; Search time 4249 Seconds

(without alignments)
9890.081 Million cell updates/sec

Title: US-10-642-531-1

Perfect score: 1104
Sequence: 1 atcgscgctgagaagctag.....cgcagcgagccttcgcatga 1104

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST1:*
1: gb_ests1:*
2: gb_ests2:*
3: gb_hncc:*
4: gb_ests3:*
5: gb_ests4:*
6: gb_ests5:*
7: gb_ests6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	878.4	79.6	1329	3	CNS0A61R
2	878.4	79.6	1360	3	CNS0A4TA
3	876.8	79.4	1397	3	CNS0A6MN
4	875.2	79.3	1350	3	CNS0A6XJ
5	874.4	79.2	1386	3	CNS0A6R4
6	531.6	48.2	793	6	CD827897
7	510	46.2	1095	9	CL961615
8	501.2	45.4	1546	3	AY110322
9	488	44.2	1438	3	AY109472
10	481.6	43.6	807	5	BX839689
11	465.2	42.1	1098	6	CL973570
12	455.4	41.2	554	6	CD831928
13	450.2	40.8	784	5	BX839466
14	441.6	40.0	1090	7	CK208011
15	409	37.0	669	7	CN892416
16	408	37.0	780	7	CF209870
17	405	36.7	807	6	CB673677
18	401	36.3	730	7	CN185956
19	398.4	35.4	851	6	CB652358
20	390.8	35.3	671	1	AV783403
21	390	35.3	841	7	CF437749
22	385.8	34.9	782	7	CF201031
23	379	34.3	774	6	CB674701
24	375	34.0	833	7	CO103634

25	371.8	33.7	730	4	BI971303
26	367.4	33.3	686	7	CN876886
27	362.6	32.8	682	7	CO481703
28	362	32.8	653	7	CN915299
29	359.2	32.5	616	7	CN907157
30	358.2	32.4	1023	4	BG837999
31	356.4	32.3	628	7	CN918862
32	355.4	32.2	815	7	CN185955
33	351.4	31.8	674	6	CB008117
34	350.6	31.8	442	1	AV557206
35	350.4	31.7	621	7	CN919108
36	350.4	31.7	785	4	BJ570152
37	350	31.7	585	7	CN896098
38	342	31.0	612	2	BE346718
39	341.6	30.9	571	2	BE803392
40	341.6	30.9	585	4	BI426713
41	341.6	30.9	585	4	BI426722
42	341.6	30.9	666	7	CN907413
43	341.4	30.9	575	6	CA936210
44	338.2	30.6	592	6	CB257858
45	336.4	30.5	706	4	BI309456

ALIGNMENTS

RESULT 1

CNS0A61R
LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

1 (bases 1 to 1329)
Unpublished
Genoscope.
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqret@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castellani V., Aury J.M., Jallion O., Wincker P., Clepet C., Schachter V., Weissenbach J., Salanoubat M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Barque_Projet_EF/Full_length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
location/Qualifiers
1. 1329
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"

gene /clone="GSLT.S392G08"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
complement(1..1329)
/gene="At3g06483"

Query Match 79.6%; Score 878.4; DB 3; Length 1329;
Best Local Similarity 87.9%; Pred. No. 2.5e-259;
Matches 970; Conservative 0; Mismatches 131; Indels 3; Gaps 1;

```

Oy 1 ATGGCGGTGAAGAGGCTGAGAGATGTTTTCGAAGAGCTTGATCGAGAGCTTTCACGA 60
Db 69 ATGGCGGTGAAGAGAGGCTGCGAAATGTTCCGAAAGATTGATGGAAGATGTTTCCAAA 128
Oy 61 TGGGATGATGATGAGAGAGCGGCGTGAAGCTCAGATCATCATGATGATGATGATGAT 120
Db 129 TGGGATGATGATGAGAGAGCGGCGTGAAGCTTGAATCATGATGATGATGATGATGAT 188
Oy 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAAGAGAGCTTCCGATTCG 180
Db 189 CCTACTGAGAGGAATCTTTTGAATTTCTGCTAGTTTTCATTAAGAGCTTCCGATTCG 248
Oy 181 ATCGGAGAGCGTGCATGAGTGAATCTGAGAGCGCTCTTAAGCGCTTCTGAGAAAAC 240
Db 249 GTGCGCAGAGAGCGATGAACTCCAGACGCTTCTTATGCTCTCTGATTAACCTGCC 308
Oy 241 GTCTTGAAGTAAAGATTTGGATGTGAGTCAATTCAGAGACATGAGAGCGCTTCCGAG 300
Db 309 GTTTGAAAGTGCAGGAGTTGGTATTTGGAATCTTTCAGAGACATGAGAGCATTTCCG 368
Oy 301 ATCAAGGATATCTGCTGATGAGAGAGATTCACACAGATGATCAAGCGCTGTTAAAG 360
Db 369 ATTAAGGATTCGGGAGAGAGAGATTTCACTCAGATGATTAAGCGCTGTTAAAG 428
Oy 361 CACAACAAGTGTGTTCCCATGATGAGCTCTGGTGAACCGCTGAAGAGAGATGAA 420
Db 429 CATACATATGCTGCTCCATATGAGCTTTGGGTGATGATGATGATGATGATGATGAT 487
Oy 421 CTTCAGAAAAGCTTGATGATGATTCATCAATTTCTGATGATGATGATGATGATGAT 480
Db 488 --TTCTGGAATCTTGATGATGATTCATCAATTTCTGATGATGATGATGATGATGAT 545
Oy 481 GGGATCGGTATGCTTATCGGCGACAGATGATGATGATGATGATGATGATGATGATGAT 540
Db 546 GGGATCGGTATGCTTATCGGCGACAGATGATGATGATGATGATGATGATGATGATGAT 605
Oy 541 ACAGTGGATTAACATACACACCAAGATGCTCTATGAGAGGAGAGAGATGATGATG 600
Db 606 ACAGTGGATTAACATACACCAAGATGCTCTATGAGAGGATGATGATGATGATGAT 665
Oy 601 GATGCAAGTTCGATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATG 660
Db 666 GATGCTCGTCAATTTTTCAGAGATGATGATGATGATGATGATGATGATGATGATG 725
Oy 661 GATCAGAGTTTCACTTTCCGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 726 GATCCAGATTTCACCTTTCCGATGATGATGATGATGATGATGATGATGATGATG 785
Oy 721 GTCAAGAACTCTCTCCGCTGCTCAAGAGCGGTTTGTGACTCTGATAGGTTGACCA 780
Db 786 GTCAAGAACTCTCTACGCTGCTCAAGAGCGATTTGTGACTCTGATAGGTTGACCA 845
Oy 781 CCAATCGGTATCAATTTGTTGCTGATGATGATGATGATGATGATGATGATGATGAT 840
Db 846 CCAATCGGTATCAATTTGTTGCTGATGATGATGATGATGATGATGATGATGATGAT 905
Oy 841 GGTGAGATTAACGAGAGCGGCTCTCCCTAAATATTCATTAAGCTTCAAGAGCTGCA 900
Db 906 GGTGAGATTAACGAGAGCGGCTCTCCAGAAATTTACCTATTTTCAAGAGCTGCA 965
Oy 901 AGAAACCCACTTGAAGAAGATGATGATGATGATGATGATGATGATGATGATGATG 960

```

```

Db 966 AGAAACCCGCTTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1025
Oy 961 GGTATGATTAAGTGTGCTGCTATTAAGCTTGTATGAGCTTGTATGATGATGATGAT 1020
Db 1026 GGTATGATTAAGTGTGCTTCCAAATTAAGCTTGTATGATGATGATGATGATGATG 1085
Oy 1021 CAGATCATTAATCAATGAGAGATACGAGACTGATGATGATGATGATGATGATGAT 1080
Db 1086 CAGATCATTAATCAATGAGAGATACGAGACTGATGATGATGATGATGATGATGAT 1145
Oy 1081 GATTCGAGAGAGCTTGTCCATGA 1104
Db 1146 GATTCGAGAGAGCTTATCACTGA 1169

```

RESULT 2
CNS044TA
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLT.S122A07 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
BX823217
ACCESSION
BX823217.1 GI:42462905
VERSION
HTC; GSLT cDNA.
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1360)

REFERENCE
AUTHORS
Menard M., Cruaud C., Quetier F., Scarpelli C., Schachter V.,
Temple G., Caboche M., Weissenbach J., and Salanoubat M.
Whole genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1360)
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT
- Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelji
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/gb/gb?source=Arabidopsis.
location/Qualifiers
1..1360

FEATURES
source
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLT.S122A07"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
complement(1..1360)
/gene="At3g06483"

ORIGIN
Query Match 79.6%; Score 878.4; DB 3; Length 1360;
Best Local Similarity 87.9%; Pred. No. 2.5e-259;
Matches 970; Conservative 0; Mismatches 131; Indels 3; Gaps 1;

QY 1 ATGGCGGTGAAGAGGCTGAGAGATGTTTTGGAAGAGCTTGATGAGAGCGTTACAGA 60
 Db 96 ATGGCAGTGAAGAGAGCGCTGGGAATGTTCCGAAGAGTTGATGGAAGATGTTACAAA 155
 QY 61 TGGGGATGATGAAGCAGAGCGGGCGTGAGCTCAGGTAATGATGAGAGTTGGTTCCACT 120
 Db 156 TGGGGTTGCAAGAGCAACCGGATGTAAGCTTGTGATACAGATGAGAGTTGGTTCCAAA 215
 QY 121 CCCCATGAGAGAAACCTTCTGATCTCGGCGCAGTTTCTTCAAGAGAGCTCCGATTCGG 180
 Db 216 CCTACTGAGAGAAATCTTTTCAATTTCTGCTCAGTTTTCATTAAGAGAGCTTCCATTCGC 275
 QY 181 ATGCGAGGCGGTGCGATGAACTCGAGACGCTGCTTAATGAGCTCTGAGAAAAGCTGCC 240
 Db 276 GTCCCGAGAGAGCGATGAACTCGAAGCGCTTCTTAATGAGCTCTGATTAAGAGAGCTGCC 335
 QY 241 GTCTTGAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 Db 336 GTTGTGAAGGTGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 395
 QY 301 ATCAAGGATACCTGCTGATGAGAGAGTTCAACAGATGATCAAGAGGCTTAAAGTAGG 360
 Db 396 ATTAAAGATTCGGGTGACAGAGAAAGATTTCACTAGATGATTAAGAGCTGCAAGATAGG 455
 QY 361 CACAACAACGTGTTCCATGATGAGCTCTGGGTGTAACAGCTGAAGAAAGAGATGAA 420
 Db 456 CATAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 514
 QY 421 CTCTACAGAAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 Db 515 -TTCTGAAATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 572
 QY 481 GGGATCCGTAATGCTTATCGGGCAGCATGTTGATGATGATGATGATGATGATGATGATG 540
 Db 573 GGGATCCGTAATGCTTATCGGGCAGCATGTTGATGATGATGATGATGATGATGATGATG 632
 QY 541 ACAAGTGGTATACATACACACCAAGATGTTCTCTATGAGAGTGGCAGAAATGCTAGTAA 600
 Db 633 ACAAGTGGTATACATACACCAAGATGTTCTCTATGAGAGTGGCAGAAATGCTAGTAA 692
 QY 601 GATGCAAGTGCATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGATG 660
 Db 693 GATGCTCGTCAATTTGTTTCCGAGATGATGATGATGATGATGATGATGATGATGATG 752
 QY 661 GATCAGAGTTCACTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
 Db 753 GATCAGAGTTCACTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 812
 QY 721 GTCAGAACTCTCTCCGCTGCTTCCAGAGCGGTTTGTGATCTGATAGGCTTGACCA 780
 Db 813 GTCAGAACTCTCTAGCTGCTTCCAGAGCGGATTTGTGATCTGATAGGCTTGACCA 872
 QY 781 CCAATCCGTAATGTTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 840
 Db 873 CCAATCCGTAATGTTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 932
 QY 841 GGTGAGGTATACGAGAGGCTCTCCCTAAATATTAATTAATTAATTAATTAATTAATTA 900
 Db 933 GGTGAGGTATACGAGAGGCTCTCCCTAAATATTAATTAATTAATTAATTAATTAATTA 992
 QY 901 AGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
 Db 993 AGAAACCACTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1052
 QY 961 GGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 Db 1053 GGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1112
 QY 1021 CAGATCATTCATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 Db 1113 CAGATCATTCATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1172
 QY 1081 GACTCGAGAGAGCTTTTGCATGA 1104

Db 1173 GATTCGCAAGAGCTTTACCTGA 1196
 RESULT 3
 CDS0A6M
 LOCUS
 DEFINITION
 CDS0A6M 1397 bp mRNA linear HTC 06-FEB-2004
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 GSLT1S88ZG07 of Adult vegetative tissue of strain col-0 of
 Arabidopsis thaliana (thale cress).
 BX823886
 ACCESSION
 BX823886.1 GI:42465058
 VERSION
 HTC; GSLT cDNA.
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 1397)
 Castell, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C.,
 Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
 Temple, G., Caboche, M., Weisenbach, J., and Salanoubat, M.
 Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 Unpublished
 2 (bases 1 to 1397)
 JOURNAL
 Genoscope.
 Direct Submission
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT
 The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castell
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weisenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information Center for Protein
 Sequences) . 5 prime and 3 prime are assembled with Phrap.
 http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_Full
 length
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
 FEATURES
 source
 1..1397
 location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSLT1S88ZG07"
 /tissue_type="Adult vegetative tissue"
 /plasmid="pCMVSPORT 6"
 complement(1..1397)
 /gene="At3g06483"
 ORIGIN
 Query Match 79.4%; Score 876.8; DB 3; Length 1397;
 Best Local Similarity 87.8%; Pred. No. 7.8e-259;
 Matches 969; Conservative 0; Mismatches 132; Indels 3; Gaps 1;
 Db 161 ATGGCAGTGAAGAGAGCTGCGAAATGTTCCGAAGAGTTGATGAGAGTGTTCACAA 220
 QY 61 TGGGGATGATGAAGCAGAGCGGCGTGAGCTCAGGTACATGATGAGAGTTCCGTTCACT 120
 Db 221 TGGGGTTGATGAAGCAACCGGATGTAAGCTTAAGATGATGATGATGATGATGATGATG 280
 QY 121 CCACATGAGAGAAACCTTCTGATCTCGGCGCACTTTCTTCAAGAGAGCTTCGATTCGG 180
 Db 281 CTTACTGAGAGAGAAATCTTTTGAATTTCTGCTCAGTTTTCATTAAGAGAGCTTCGATTCGC 340

QY 181 ATGCGAGGCGTGCATCGAAGCTCGAGAGCTGCTTATGCGCTCTGAGAACTCC 240
 DB 341 GTGCGCAGAGAGCCATCCAGACGCTTCTTATGCTCTCTGATTAACCTCC 400
 QY 241 GTCTTGAAGTAAAGATTGTATGTGAGTATTCAGGACATGAGACCTTCTGAG 300
 DB 401 GTTTTGAAGTCCGGGATTTGTATTTGGAATCTTTCAGGACATGAGACATTTCTAG 460
 QY 301 ATCAAGGATATCTGATGAGAAAGTTCACAGATGATCAAGCTGTTAAAGTAAAG 360
 DB 461 ATTAAGATTCGGGTGACAGAAAGATTTCACTGATGATTAAGCTGTCAAAAGTAAAG 520
 QY 361 CACAAACAAGCTGTTCCATGATGCTGTGAGTGAACAGAGTGAAGAAAGTAA 420
 DB 521 CATTAACAATGTGTTCCCATGATGCTTGGGTGTTAATCAAGCTCAAAAGATGA 579
 QY 421 CTCTTGAAGAAAGCTGATGAGATTATCATGTTTCTTATGCTGCTTCTGATTA 480
 DB 580 -TTCTGGAATCTGATGAGATTCATCATGTTCTTATGCTGCTTCTGATTA 637
 QY 481 GGGATCCGATGCTTATGCGGAGCATGTTGATGCTTATGCTTATGCTTATGCTTAT 540
 DB 638 GGGATCCGATGCTTATGCGGAGCATGTTGATGCTTATGCTTATGCTTATGCTTAT 697
 QY 541 ACAGTGGTAAATACACACCAAGATGCTCTTATGAGTGGCAGAGATGCTAGTAA 600
 DB 698 ACAGTGGTAAATACACCAAGATGCTCTTATGAGTGGCAGAGATGCTAGTAA 757
 QY 601 GATGCAAGTGAATTTGTTTCAAGAGATGCTTCTGCTGCGGATTAATATATATG 660
 DB 758 GATGCTGCTGATTTGTTTCAAGAGATGCTTCTGCTGCGGATTAATATATATG 817
 QY 661 GATCAAGTTCATTTTCTGATGCTTCAAGAGATGCTTCTGCTGCGGATTAATATAT 720
 DB 818 GATCCAGTTCATTTTCTGATGCTTCAAGAGATGCTTCTGCTGCGGATTAATATAT 877
 QY 721 GTCAAGATCTCTCCGCTGCTTCAAGAGATGCTTCTGCTGCGGATTAATATATAT 780
 DB 878 GTCAAGATCTCTCCGCTGCTTCAAGAGATGCTTCTGCTGCGGATTAATATATAT 937
 QY 781 CCAATCCGATCATTTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 840
 DB 938 CCAATCCGATCATTTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 997
 QY 841 GGTGAGATATCCGAGAGAGGCTCTCCCTTAAATATATCACTTACAGACATGCA 900
 DB 998 GGTGAGATATCCGAGAGAGGCTCTCCCTTAAATATATCACTTACAGACATGCA 1057
 QY 901 AGAAACCCACTTGAAGAGATGAGTGAACCGCTGATGTTCCCTGATATGCTG 960
 DB 1058 AGAAACCCACTTGAAGAGATGAGTGAACCGCTGATATGCTGATATGCTGAT 1117
 QY 961 GGTATGATATGCTGCTTATGATGCTTATGATGCTTATGATGCTTATGATGCT 1020
 DB 1118 GGTATGATATGCTGCTTATGATGCTTATGATGCTTATGATGCTTATGATGCT 1177
 QY 1021 CAGATCATATCCATGAGAGATGAGGATGATGATGATGATGATGATGATGATGAT 1080
 DB 1178 CAGATCATATCCATGAGAGATGAGGATGATGATGATGATGATGATGATGATGAT 1237
 QY 1081 GACTCGAGAGAGCTTGGCATGA 1104
 DB 1238 GATTGCGAAGAGCTTTACCTGA 1261

RESULT 4
 CDS0A6XJ 1350 bp mRNA linear HNC 06-FEB-2004
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 DEFINITION GSTRPH0204 of Hormone Treated Callus of strain col-0 of
 Arabidopsis thaliana (thale cress).
 ACCESSION BX825037

VERSION
 EX825037.1 GI:42465807
 HNC; GSTR CDNA.
 KEYWORDS
 Arabidopsis thaliana (thale cress)
 SOURCE
 Arabidopsis thaliana
 ORGANISM
 Arabidopsis thaliana
 REFERENCE
 Castell, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C.,
 Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
 Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
 Whole Genome Sequence Comparisons and Full-length cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 JOURNAL
 Unpublished
 REFERENCE
 2 (bases 1 to 1350)
 Genoscope.
 DIRECT SUBMISSION
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr
 JOURNAL
 Web : www.genoscope.cns.fr)
 COMMENT
 The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castell
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 URV INRA, Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Piprap.
 http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
 length
 http://www.genoscope.cns.fr/cgi-bin/gdb/gdb?source=Arabidopsis.
 FEATURES
 location/Qualifiers
 1..1350
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="COL-0"
 /db_xref="taxon:3702"
 /clone="GSTRPH0204"
 /tissue_type="Hormone Treated Callus"
 /plasmid="pCWSPORT 6"
 complement(1..1350)
 /gene="At3g06483"
 ORIGIN
 Query Match 79.3%; Score 875.2; DB 3; Length 1350;
 Best Local Similarity 87.7%; Pred. No. 2.4e-258;
 Matches 968; Conservative 0; Mismatches 133; Indels 3; Gaps 1;
 QY 1 ATGCGGTGAAGAGGCTAGAGATGTTTCAAGAGCTTGAACGAGACGTTCAAGA 60
 DB 144 ATGCGAGTAAAGAAAGCTGCGAAATGTTCCGAGAGTTGATGCAAGATGTTCAAAA 203
 QY 61 TGGGATGATGAAGACAGACGCGGTGAGCTCAGTACATGATGAGAGTTGCTTCACT 120
 DB 204 TGGGATGATGAAGACAGACGCGGTGAGCTTATGATATCATGATGAGAGTTGCTTCA 263
 QY 121 CCGACTGAGAGAACTTGTGATCTCGCGCAGATTTCTTCAAGAGAGCTTCCGATTCG 180
 DB 264 CTACTGAGAGAACTTGTGATTTGCTCAAGTTTGTGATGAAGAGCTTCGATTCGC 323
 QY 181 ATGCGAGGCGTGCATGCAACTGAGACGCTGATAGGCTCTCTGAGAAACCTGCC 240
 DB 324 GTGCGCAGAGAGAGATGCAACTCAGACCTTCTTATGATGATGATGATGATGATGAT 383
 QY 241 GTCTTGAAGTAAAGATTGTATGTGAGTATTCAGGACATGAGAGGCTTCTGAG 300
 DB 384 GTTTTGAAGTCCGGGATTTGTATTTGGAATCTTTCAGGACATGAGACATTTCTGAG 443
 QY 301 ATCAAGATCTGCTGATGAGAAAGTTCACAGATGATCAAGCTGTTAAAGTAAAG 360

Db 444 ATTAAGGATTGGGTGACGAGAAAGATTTCATCTAGATGATTAAGGCTGTCAAGTAAAG 503
 Qy 361 CAGAACAACTGGTTCCTCCATGATGCTCTGGGTGTGAACAGCTGGAAGAAAGAAATAA 420
 Db 504 CATACAAATGTGTCTCCATGATGAGCTTGTGGTGTATATAGCTCAAGAAAGAAATGA- 562
 Qy 421 CTTCAGAAAAGCTGTAGATGATTCATGATTTCTTATGCTGTCTACTGTCGTATA 480
 Db 563 --TTCTGGAATCTTGATGATGATTCATGATTTCTTATGCTGTCTACTGTCGTATA 620
 Qy 481 GGGATCCGATGCTTATCGGACAGCATGTTGAGTTCATAATCCAAACCAACCTTAC 540
 Db 621 GGGATCCGATGCTTATCGGACAGCATGTTGAGTTCATAATCCAAACCAACCTTAC 680
 Qy 541 ACAGTGGTTCATACACCAAGATGCTCTTATGAGGTGCAAGAAATGCTATGAA 600
 Db 681 ACAGTGGTTCATACACCAAGATGCTCTTATGAGGTGCAAGAAATGCTATGAA 740
 Qy 601 GATGCAAGTTCATGTTGTTCAAGAGTATGTTCTCTCCGAGATTAACATATATGAC 660
 Db 741 GATGCTCGTCAATTTGTTCCAGAGTACGTTCTGACCCGAAATTAACATATATGAC 800
 Qy 661 GATCCAAATTCATTTCTCCGATGTTCCGACCCCATTTGATGTTATGATGATTA 720
 Db 801 GATCCCAATTCATTTCTCCGATGTTCCGACCCCATTTGATGTTATGATGATTA 860
 Qy 721 GTCAGAACTCTCTCCGATGTTCCGACCCCATTTGATGTTATGATGATTA 780
 Db 861 GTCAGAACTCTCTCTCCGATGTTCCGACCCCATTTGATGTTATGATGATTA 920
 Qy 781 CCAATCCGATTCATTTGTTGCTGATGATGATGATGATGATGATGATGATGATGAT 840
 Db 921 CCAATCCGATTCATTTGTTGCTGATGATGATGATGATGATGATGATGATGATGAT 980
 Qy 841 GGTGAGATGATTCAGAGACCGGTCTCCCTAAATATTCATCTTACAGACTGCA 900
 Db 981 GGTGAGATGATTCAGAGACCGGTCTCCCTAAATATTCATCTTACAGACTGCA 1040
 Qy 901 AGAAGCCCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 Db 1041 AGAAGCCCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1100
 Qy 961 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 Db 1101 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1160
 Qy 1021 CAGATCATATTCATGAGAGATACGGGATGATGATGATGATGATGATGATGATGAT 1080
 Db 1161 CAGATCATATTCATGAGAGATACGGGATGATGATGATGATGATGATGATGATGAT 1220
 Qy 1081 GACTCGCAGAGAGCTTTGGCATGA 1104
 Db 1221 GATTGCGAAGAGCTTTTACCTGA 1244

RESULT 5
 LOCUS CDS0A6R4 1386 bp mRNA linear HTC 06-FEB-2004
 DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 GSTRGH56ZC08 of Hormone Treated Callus of strain col-0 of
 Arabidopsis thaliana (thale cress).
 ACCESSION BX824559.1 GI:42465670
 VERSION 1
 KEYWORDS HTC; GSTR cDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Bacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1386)
 Castelli, V., Aury, J.M., Jallion, O., Winkler, P., Clepet, C.,
 Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
 Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.

TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1386)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage ;
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli
 V., Aury J.M., Jallion O., Winkler P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.
 http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
 length
 http://www.genoscope.cns.fr/cgi-bin/g9b/g9b?source=Arabidopsis.
 FEATURES
 source 1.1386
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSTRGH56ZC08"
 /issue_type="Hormone Treated Callus"
 /plasmid="pCMVSPORT 6"
 complement(1..1386)
 /gene="At3g06483"
 ORIGIN
 Query Match 79.2%; Score 874.4; DB 3; Length 1386;
 Best Local Similarity 87.8%; Pred. No. 4,3e-258;
 Matches 966; Conservative 0; Mismatches 131; Indels 3; Gaps 1;
 Qy 1 ATGCGGTGAGAGAGCTGACGATGTTTTCGAAGAGCTTTCAGAGAGCTTTCACAGA 60
 Db 134 ATGCGAGTTAAGAAACCTGCGAAATGTTCCCAAGAGTTGATTCAGAAATGTTCCAA 193
 Qy 61 TGGGATGATGAGAGAGACGCGGCTGAGCTCAGGATACATGATGAGTTCCGTTCACT 120
 Db 194 TGGGATGATGAGAGAGACGCGGCTGAGCTCAGGATACATGATGAGTTCCGTTCAAA 253
 Qy 121 CCCACTGAGAGAAACCTTCGATCTCGGCGAGTTTCTTCACAGAGAGCTTCCGATTCGG 180
 Db 254 CTTACTGAGAGAAATCTTTTGTATTTCTGCTCACTTTTTCATAGAGAGCTTCCGATTCGC 313
 Qy 181 ATGCGAGCGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 Db 314 GTCGCGAGAGAGCGATTCAGAACTCCAGACGCTCTTATGATGATGATGATGATGATGAT 373
 Qy 241 GTCCTAAGGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 Db 374 GTTTTAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 433
 Qy 301 ATCAAGATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 Db 434 ATTAAGATTCGGGTGACGAGAAAGATTTCACTCATATGATTAAGGCTGTCAAAAGTAAG 493
 Qy 361 CACACAAAGTGTTCCTCATGATGCTCTGGGTGTAAACAGCTGAAGAAAGAAATGAAA 420
 Db 494 CATTAACAATGTGTCTCCATGATGCTTGGGTGTAAACAGCTGAAGAAAGAAATGAAA- 552
 Qy 421 CTTCAGAAAAGCTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 Db 553 --TTCTGGAATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 610
 Qy 481 GGGATCCGATGCTTATCGGACAGCATGTTGAGTTCATATCCAAACCAACCATTTAC 540

Db 611 GGGATCCGATGCTTATTTGGGAGCAAGTTGATGTCATATTCATCAATCCACCGCTTCAAT
Qy 541 ACGTGGGTATACATACACCAAGATGTCCTATGAGAGTGGCAAGAACTGATGAA
Db 671 ACGTGGGTATATACACCAAGATGTCCTATGAGAGTGGCAAGAACTGATGAA
Qy 601 GATGCAAGTTCATTTGTTTTCAGAGATAGTTCGCTCGGAGATTAACATATATGAC
Db 731 GATGCTCGGTCAATTTGTTTCGAGAGTACGTTCTGACCGGAAATTAACATATATGAC
Qy 661 GATCCAAAGTTCACTTTCCGATGTTCCGACCCATTTGATCTTATGATGATGATTA
Db 791 GATCCAAAGTTCACTTTCCGATGTTCCGACCCATTTGATCTTATGATGATGATTA
Qy 721 GTCAAGAACTCTCCGATGTTCCGACCCATTTGATCTTATGATGATGATTA
Db 851 GTCAAGAACTCTCTGATGTTCCGACCCATTTGATCTTATGATGATGATTA
Qy 781 CCAATCCGATATGTTGCTGATGAAATCGAAGATGTTAAATTAAGTCTCAAGTAA
Db 911 CCAATCCGATATGTTGCTGATGAAATCGAAGATGTTAAATTAAGTCTCAAGTAA
Qy 841 GGTGAGGTATACCGAAGAGCGGTCTCCCTAAATATTCATTAATCTCTACAGACCTGA
Db 971 GGTGAGGTATACCGAAGAGCGGTCTCCCTAAATATTCATTAATCTCTACAGACCTGA
Qy 901 AGAAACCCACTTGAAGAGATGAGACTTGGAAACCGGTGATGTTCCCGTACATAGCT
Db 1031 AGAAACCCACTTGAAGAGATGAGACTTGGAAACCGGTGATGTTCCCGTACATAGCT
Qy 961 GGTATGTTATGTTGCTGCTATGATGCTGTTGATGCTGTTGATGCTGTTGATGCT
Db 1091 GGTATGTTATGTTGCTGCTATGATGCTGTTGATGCTGTTGATGCTGTTGATGCT
Qy 1021 CAGATCATATCATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT
Db 1151 CAGATCATATCATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT
Qy 1081 GACTTCGACGAGACCTTTGCC 1100
Db 1211 GATTTCGACGAGACCTTTACC 1230

RESULT 6
CD827897 793 bp mRNA linear EST 10-JUL-2003
LOCUS BN25.069A22F020124 BN25 Brassica napus cDNA clone BN25069A22, mRNA
DEFINITION sequence.
ACCESSION CD827897
VERSION CD827897.1 GI:32509837
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 793)
AUTHORS Genoplane, a major partnership french program in plant genomics
TITLE Genoplane (2003)
JOURNAL Unpublished
COMMENT Contact: Genoplane
Genoplane
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplane' (<http://www.genoplane.com>
and <http://genoplane-info.infobiogen.fr>).
FEATURES
source 1..793
/organism="Brassica napus"
/mol_type="mRNA"

ORIGIN
/cultivar="Jet Neuf"
/db_xref="taxon:3708"
/clone="BN25069A22"
/issue_type="seed"
/clone_id="BN25"

Query Match 48.2%; Score 531.6; DB 6; Length 793;
Best Local Similarity 88.6%; Pred. No. 2.8e-152;
Matches 576; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 24 GATGTTTGAAGAGTTGATGAGAGACCTTCAAGATGAGATGAGATGAGATGAGATGAGAT
Db 144 GATGTTTGAAGAGTTGATGAGAGACCTTCAAGATGAGATGAGATGAGATGAGATGAGAT
Qy 84 CGTGAGCTCAGTATCATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT
Db 204 CGTGAGCTCAGTATCATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT
Qy 144 CTGAGCTCAGTATCATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT
Db 264 CTGAGCTCAGTATCATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT
Qy 204 CGAGAGCTGCTTATGAGCTGCTGAGAGAACTGCTGAGAGAACTGCTGAGAGAACTGCTGAG
Db 324 CGAGAGCTGCTTATGAGCTGCTGAGAGAACTGCTGAGAGAACTGCTGAGAGAACTGCTGAG
Qy 264 TGTGAGTCAATTCAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT
Db 384 TGTGAGTCAATTCAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT
Qy 324 AGAGTTTCAAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT
Db 444 AGAGTTTCAAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT
Qy 384 GGCCTGAGTGAAGCAAGTGAAGAAAGATGAAGAACTGACGAAAGCTTGAATGAGAT
Db 504 GGCCTGAGTGAAGCAAGTGAAGAAAGATGAAGAACTGACGAAAGCTTGAATGAGAT
Qy 444 TCATGAGTTTCTGATGCTTCTGATGCTTCTGATGCTTCTGATGCTTCTGATGCTTCTGAT
Db 564 TCATGAGTTTCTGATGCTTCTGATGCTTCTGATGCTTCTGATGCTTCTGATGCTTCTGAT
Qy 504 GCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db 624 GCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy 564 GATGCTCTCTATGAGAGTGGCAAGAAATGCTAGTGAAGATGCAAGTCTGATTTGTTTCA
Db 684 GATGCTCTCTATGAGAGTGGCAAGAAATGCTAGTGAAGATGCAAGTCTGATTTGTTTCA
Qy 624 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db 744 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT

RESULT 7
CU961615 1095 bp DNA linear GSS 21-SEP-2004
LOCUS OsiRCC006652 Oryza sativa Expressed Library Oryza sativa (indica
DEFINITION cultivar-group) genomic, genomic survey sequence.
ACCESSION CU961615
VERSION CU961615.1 GI:52377973
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; liliopsida; Poales; Poaceae;
Euphorbiaceae; Oryzaeae; Oryza.
REFERENCE 1 (bases 1 to 1095)
AUTHORS Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. W., and Wang, J.

TITLE An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatics
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES

source 1. .1095
Location/Qualifiers
/organism="Oryza sativa (indica cultivar-group)"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match 46.2%; Score 510; DB 9; Length 1095;
Best Local Similarity 68.4%; Pred. No. 1.5e-145;
Matches 746; Conservative 0; Mismatches 320; Indels 24; Gaps 2;

Oy 22 GAGATGTTTCAAGAGCTTGAATCAGAGAGCTTCAAGATGGGATGATGAAGCAGACG 81
Db 13 GACCCGCGCGGAGGCGGCGGAGAGAGTGGCGGAGGAGATGAGCAGACG 72
Oy 82 GCGGTGAGCTTCAGATGATGATGAGTGGTTTCACTCCACTGAGAGAACTTTCTG 141
Db 73 GCGGTGAGCTTCAGATGATGATGAGTGGTTTCACTCCACTGAGAGAACTTTCTG 132
Oy 142 ATCTCGCGGAGTCTTCAAGAGAGCTTCCATTCGATGCGAGCGCTGATGCGA 201
Db 133 CGGTGCGGAGTCTTCAAGAGAGCTTCCATTCGATGCGAGCGCTGATGCGA 192
Oy 202 CTCGAGAGCGTGCCTTATGAGCTTCTCTGAGAGAACTGCGCTTGAAGTAAAGATTGG 261
Db 193 CTCGATCCTCTCCCTTCCGCTCTCTCCAGAGCGGCTCTCTCAAGATGAGATTGG 252
Oy 262 TATGTGAGTCAATTCAGAGAGATGAGAGCGTTCTTCTGATGATCAAGATCTGATGAG 321
Db 253 TATTTGAGACTATTCGCGAGCTTACCGTGTCTCCGAGGTGAGAGACCGGATGATGAG 312
Oy 322 AAGAGTTCACAGAGATGATCAAGAGCTTAAAGTAAAGAGAGAGAGAGAGAGAG 381
Db 313 CTTCATTCACAGAGATGATCAAGAGCTTAAAGTAAAGAGAGAGAGAGAGAG 372
Oy 382 ATGGCTCTGGGTGAGAGAGCTGAGAGAG-----GAATGAAGCTTACGAA 429
Db 373 ATGGCTCTGGGTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 432
Oy 430 AAGCTTGATGATGATCAAGAGCTTCTGATGCTTCTGATGCTTCTGATGAGAGAG 489
Db 433 GCGATGATGATGATCAAGAGCTTCTGATGCTTCTGATGCTTCTGATGAGAGAG 492
Oy 490 ATGCTTATCGGAGAGATGATGAGAGCTTCAATTCAGAGAGAGAGAGAGAGAGAG 549
Db 493 ATGCTTATCGGAGAGATGATGAGAGCTTCAATTCAGAGAGAGAGAGAGAGAG 552
Oy 550 TACATACACAGAGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 609
Db 553 CTGATGAGAGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 612
Oy 610 TCGATTTGTTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 669
Db 613 GCTATTTGATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 672
Oy 670 TTGATTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 729
Db 673 TTGATTTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 732

Oy 730 TCTCTCGTCTGCTCCAGAGAGCGTTGTTGACTCTGATGAGGTTGACCAACATCCGT 789
Db 733 TCCCTCGTCTGCTCCAGAGAGCGTTGTTGACTCTGATGAGGTTGACCAACATCCGT 792
Oy 790 ATCATTTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 849
Db 793 ATCATTTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 852
Oy 850 ATACCGAGAGAGGCTCTCCAGAGAGCGTTGTTGACTCTGATGAGGTTGACCAAC 909
Db 853 ATACCGAGAGAGGCTCTCCAGAGAGCGTTGTTGACTCTGATGAGGTTGACCAAC 912
Oy 910 CTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 969
Db 913 CTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
Oy 970 TATGTTCTGCTTATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1029
Db 961 TATGAGATCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Oy 1030 TCCATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1089
Db 1021 TCTATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Oy 1090 GAGCCTTTGC 1099
Db 1081 GAGCCTTTGC 1090

RESULT 8
AY110322 1546 bp mRNA linear HTC 17-OCT-2002
LOCUS Zea mays CL1703_1 mRNA sequence.
DEFINITION AY110322
ACCESSION AY110322
VERSION AY110322.1 GI:21214656
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1546)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whiteside, M.S.,
Arthur, L.W., Hanefey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/Dupont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 1546)
Coe, E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schubert, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
Location/Qualifiers
1. .1546
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/Dupont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by Dupont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

FEATURES

source

D	b		237	GCTCGA	CCTCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCAAAGCCGGCATTCTCAAGGTGA	296
O	y		255	AGATTGTA	TGTGTAGATCATTCAGGACATGAGAAGCGTTTCTGAGATCAAGATACTGC	314
D	b		297	AGATTGTA	TGTGTAGATCATTCGGTAAATCCGGTCTCTTCCAGAGTGAGAACAGAA	356
O	y		315	TGATGAAA	AGAGTTCAACAAGATGATCAAAGCTGTAAAGTAAGGCACAAACGTGCT	374
D	b		357	AGATAGCT	CGGCTTCAACCAGATGATCAAAATGATCAGGGTAGCACACCAATGTGCT	416
O	y		375	TCCCATGA	TGAGCTGTGGGTGTAACCAAGCTGAAGAAAG-----GAATGAAACT	422
D	b		417	GCCTCGAT	TGCAATTGGGTGTGCAGACAGCTGAAGAAAGACTTGGGGGCTCCAAGCGTT	476
O	y		423	CTACGAAA	AGCTTGAATGATTCATCAGTTTCTTGATCGCTTACTGTCTCTGATPAG	482
D	b		477	NNNNNN	TGGAAATCCATGAGATCCATCAGTTCTTTCGCCCTTCAATGTCAAAGATTGG	536
O	y		483	GATCCGTA	TGCTTAATCCGGCAGACATGTTAGTTGCAATATCCAAACCCACACTTCACAC	542
D	b		537	GATCCGAT	CTGCTGATAGGACACACGTCGCTTTCATGACCTGATTCAGAGCTGTGCT	596
O	y		543	AGTGGTTA	CATACACACCAAGATGTTCTCTATGAGAGTGGCAAGAAATGCTAGTAGA	602
D	b		597	GATAGGGC	TATPAACANNNTGTGACCCATGACAGTTGCTCGAATGTCTAGTAGAA	656
O	y		603	TGCAAGGT	TCGATTTGTTCAAGAGATATGCTTCTGCTCCGAGATPAACATATATAGGCA	662
D	b		657	TGCAGTGT	CTAATTCAGTCGGGAGATATGATCATCTCGTAGTCGACATPATATGCTGA	716
O	y		653	TCCAGTTT	CACTTTTCCGATGTTCCGACCCCATTTGCATCTTATGTTAGTATGATTAGT	722
D	b		717	TCCGGGTT	TACATTTCCATATGTTACACCAATCTACATCTCAATGATTTTGAAGCTGCT	776
O	y		723	CAGAAC	CTCTCTCCGTGCTGTCCAAAGGCGTTTGTTGACTGTATAGGGTTGCACACC	782
D	b		777	GAAGAATC	CCCCCTCCGCGGTACAGGACGGTATATGATTCGATTAACCTTGACACTCC	836
O	y		783	AATCCGTA	TCATTTGTTGCTGATGGAATGCAATATGTTCAATPAAGTCTCAGATGAAG	842
D	b		837	AGTTGAAT	CAATATGTTGCTGATGAGACAGAGATGTAATTAATTAAGATGATGAAG	896
O	y		843	TGAGAGTA	TACCGAAGGCGTCTCCCTAAATATATTCATCTACCTCTACAGCACTGCAAG	902
D	b		897	TGTTGGA	TATCCAGAGATGAGGCTCTCAAGAAATTTCTACGTATCTCTATATAGTACAG	956
O	y		903	AAACCCAC	TTGAAGATGTGAACTTGGAAACCGCTGATGTTCCCTGACTATAGGCTGG	962
D	b		957	AAACCTCT	CTGATCGATGACA-----TAATGAAGAGTAACTATAGGCTGG	1004
O	y		963	TTATGTTA	TGATGCTATGATAGTCGTTGATGCTGCTATTTTGGTGGAGATTGCA	1022
D	b		1005	ATATGTTA	TGAGATCCCTATTAATGTCCTTTATGCTGATATTTTCGGTGGGACCTTGA	1064
O	y		1023	GATCATAT	CACTGGAAGTACGGGACTGATGCTTATCTTGCACTGTCTGCTTTGAGGA	1082
D	b		1065	GATCATCT	CTATATGGAAGTATGGAACATGATGCTTATCTTCACACCTTTACCGGCTGGAGA	1122
O	y		1083	CTCGCAGA	AGCTTTGCCATGA	1104
D	b		1125	TTCCGAGA	AAACCTTTCGGTAA	1146

RESULT 10				
LOCUS	BX839689	807 bp	mRNA	EST 11-FEB-2004
DEFINITION	BX839689 Arabidopsis thaliana Adult vegetative tissue Col-0			
ACCESSION	Arabidopsis thaliana cDNA clone GSLTUS3Z2E08 5PRIM, mRNA sequence.			
VERSION	BX839689			
KEYWORDS	BX839689.1 GI:42533772			
SOURCE	EST.			
	Arabidopsis thaliana (thale cress)			

ORGANISM	REFERENCE	TITLE	JOURNAL	COMMENT	FEATURES
Arabidopsis thaliana	Castelli V., Aury J.M., Jallion O., Mincker P., Clepet C., Menard M., Cruaud C., Querier F., Scarcelli C., Schachter V., Temple G., Caboche M., Weissenbach J. and Salanoubat M.	Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences	A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation	Unpublished (2004)	Genoscope Contact: Genoscope 2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Mincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST http://www.genoscope.cns.fr/cgi-bin/gdb/gdb?source=Arabidopsis.
Arabidopsis thaliana	1 (bases 1 to 807)				location/Qualifiers 1..807 /organism="Arabidopsis thaliana" /mol_type="mRNA" /ecotype="Col-0" /db_xref="taxon:3702" /clone="GSLRLS932E08" /cissue="Adult vegetative tissue" /clone_id="Arabidopsis thaliana Adult vegetative tissue Col-0"
Query Match	43.6%	Score 481.6;	DB 5;	Length 807;	
Best Local Similarity	84.0%	Pred. No. 7.8e-137;			
Matches 568;	Conservative 0;	Mismatches 104;	Indels 4;	Gaps 2;	
1	ATGGCGGTGAAGAGCTTACGAGATGTTTTCGAAGAGCTTGATCGAGAGCTTCACAGA	60			
132	ATGGCAGTGAAGAAAGCTCGAGATTTCCCGAAGAGTTTGATCGAAGATGTTTCACAAA	191			
61	TGGGATGTCATGAAGACGAGCGGCGTAGCGCTCAGTACATGATGAGAGTTGGTTCAC	120			
192	TGGGTTGTCATGAAGACAAACCGGTTAGCTTATGATATCATGATGAGAGTTGTTCCAAA	251			
121	CCCACTGAGAGAAACCTTTTGATCTCGGCGCAGTTTCTTCAACAGAGCTTCGATTCGG	180			
252	CTTACTAGAGGAAATCTT-TGATTTCTGCTCAGTTTTCGATTAAGAGAGCTTCGATTCGC	310			
181	ATGCGAGGCGGTGATCGAAGCTCGAGAGCGTGCCTTATGAGCTCTCGAAGAACTTCC	240			
311	GTGCGCGAGAGAGCATCGAGCTTCACAGCGCTTCTTATGCTCTCTGTGAATAACTGCC	370			
241	GTTTGAAGGTGAAGATTGGTATGATGAGATCATTCAGGACATGAGAGCGTTTCTGAG	300			
371	GTTTGAAGGTGCGGAGTTGGTATTTGGAATCTTTCAGGAGCATGAGAGCATTCCTGAG	430			
301	ATCAAGATCTGCTGATGAGAAAGATTTCACACAGATGATCAAGGCTGTTAAAGTAAAG	360			
431	ATTAAAGATTCGGGTGCGAAGAAATTTCACTCAGATGATTAAGGCTGTCAAAAGTAAAG	490			
361	CACAAACAAGTGTGTTCCATGATGAGCTCTGGGTGTAACACAGCTGAGAAAGAAATGAA	420			
491	CATACAAATGTGTTCCCATGATGAGCTTTGGGTGTTAATCAGCTCAAGATGAGAAATGA	548			
421	CTTACGAAAAGCTTGAATGATTCATCAAGTTCTTATGATGCGCTTCACTGTTCTCGTATA	480			

Db 549 -GTTCTGGAATCTTGATAGATTCACTGTTTCTTGATCTGTTTCTACTTGTCGGAATC 607
 QY 481 GGGATCCGATATGTTATCGGCGACGATGTTGATGCATATCCAAACCCACTTCAC 540
 Db 608 GGGATCCGATGTTATGTTGGGACGACGTTGATGTCATATCCATATCCACCGCTTCAT 667
 QY 541 ACAGTGGGTTACATACACCAAGAATGTCCTATGAGGAGTGGCAAGAACTCTGTAA 600
 Db 668 ACAGTGGGTTATTTCTCTCAAAAGATGTCCTATGAGGAGTGGCAAGAACTCTGTAA 727
 QY 601 GATCGAAGTGCATTTGTTTTCAGAGATATGTTCTGCTCCGAGATATTAATATGAC 660
 Db 728 GATGCTCGGTCAATTTGTTTCCGAGATACGTTCTGACCGGGGATGACATATATGTC 787
 QY 661 GATCCAAAGTTCACTT 676
 Db 788 GATCCAGTTTCCCCCTT 803

RESULT 11
 CL973570 1098 bp DNA linear GSS 21-SEP-2004
 LOCUS CL973570
 DEFINITION OsIRCC024520 Oryza sativa Expressed Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
 CL973570
 ACCESSION CL973570.1 GI:52401665
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Department of Bioinformatic
 Beijing Institute of Genomics
 Chinese Academy of Sciences, Beijing 101300, China
 Tel: 86-10-80481559
 Fax: 86-10-80488676
 Email: chenchen@genomics.org.cn
 Rice genomic sequence.
 Class: exon-trapped.
 Location/Qualifiers
 1..1098
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="genomic DNA"
 /db_xref="taxon:39946"
 /clone_lib="Oryza sativa Expressed Library"
 /note="Oryza sativa exon trapped genomic sequences"

FEATURES
 source

ORIGIN

Query Match 42.1%; Score 465.2; DB 9; Length 1098;
 Best Local Similarity 67.4%; Pred. No. 1e-131;
 Matches 728; Conservative 0; Mismatches 298; Indels 54; Gaps 3;

QY 13 AAGGCTAGCGAGATGTTTGGAGAGCTTATGAGAGCTTACAGATGGGGATGATG 72
 Db 1 ATGGCGTGGAGCGCGGCGGCGAGCGCGGTGCGAGAGAGTGGGAGTGGGGAGCATG 60
 QY 73 AAGCAGACGGGCGGTGAGCCTCAGGTACATGATGAGTTCGTTCCACTCCACTGAGAA 132
 Db 61 AAGCAGACGGGCGGTGAGCCTCAGGTACATGATGAGTTCGTTCCACTCCACTGAGAA 120
 QY 133 AACCTTCTGATCTTCGGGCGCATTTTTCACAGAGCTTCCGATTTGATGATCGAGGCGT 192
 Db 121 AACCTCTGCTCTCCGCGGCGATTTCTGAGAGAGAGCTCCCAATCGGATCGCGGCGCG 180

QY 193 GCGATCGAACTCGACAGCGTGCCTTATGCGCTCTCTGAGAAACCTGCGTCTTGAGGTA 252
 Db 181 GCGCTGAGCTCGAGTCCCTCCCTCCGCTCTCCCGAAGCCGCCATCTCTCAAGATA 240
 QY 253 A-----GAGATTGATGAGTCAATTCAGGAC 282
 Db 241 ATTCTGGGGAGTGGAGTGTGAATAGTGCAGGATGCTGATCTTTCGCGCAG 300
 QY 283 ATGAGACGCTTCTCTGAGATCAAGATATGCTGATGAGAGAAAGATTCACAGATGATC 342
 Db 301 ATGAGATCTTCTCAGAAAGTGAAGATAGAAACGACAGCTGCTTTTAGCAGATGATC 360
 QY 343 AAGCGCTTTAAAGTAAAGGACAAACAGTGTTCCTCATGATGAGCTCTGGGTGAGACAG 402
 Db 361 AAGATGATCAAGATAGCCCAATATATGTTTCCAAATATGAGCACTGGGATTCACAG 420
 QY 403 CTGAGAGAAG-----GAATGAACTCTACGAAAGCTTATGAGATTCATCAG 450
 Db 421 CTGAGAGAGAGAGATGATGCAACAGAAAGATCCCACTGATTTAGAGAGATTCACAG 480
 QY 451 TTTCTTGATCGCTTCTACTTGTCTGTATAGGATCCGATGCTTATCGGCGAGATGTT 510
 Db 481 TTTCTTGACAGATTTCTACATGTCAGAGATGGGATCCGATCTTATAGGCGAGATGTT 540
 QY 511 GAGTTCATATATCCAAACCCACTTACACAGTGGGTTACATACACCAAGATGCT 570
 Db 541 GCTTTCATGATTCCTGACAGAGAGCCGCGCTCATAGGCTTCATATATACAGAAATGTC 600
 QY 571 CCTATGAGAGTGGCAAGAAATGCTATGAGAGATGAGAGTGGATTTGTTAGAGATAT 630
 Db 601 CCTATACAGTGGCTCAAGCTCCAGTGAAGATGCCCTTCTATTTGTTAGAGAAATAT 660
 QY 631 GCTTCTGCTCGGAGATTAACATATATGAGAGATCCAGATTTCCATTTCCGATGTTCCG 690
 Db 661 GGATGAGCTCCGAGATGACATCTATGAGAGACCAATTTTACATTTCCATATGTTTCA 720
 QY 691 ACCCATTTGCAATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 750
 Db 721 TCACATCTACATCTCTATCTATTTGATGATGATGATGATGATGATGATGATGATGAT 780
 QY 751 CGGTTTGTGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 810
 Db 781 CGATATATGATATTCGATTAAGATGATGATGATGATGATGATGATGATGATGATGATG 840
 QY 811 GAAATGTTTAAATTAAGTCTCAGATGATGATGATGATGATGATGATGATGATGATGATG 870
 Db 841 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 QY 871 AAATATTCACCTTACCTTACAGACCTGCAAGAAACCACTTGAAGAAATGATGATGATG 930
 Db 901 AGAATTTTCAATATCTATATATAGACCTGCAAGAAATGATGATGATGATGATGATGATG 948
 QY 931 GGAACCGTGAATGTTCCCTGACATGATGATGATGATGATGATGATGATGATGATGATG 990
 Db 949 GATGCGCTTATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1008
 QY 991 TTGATGCTGCTATTTTGGGAGATTTGGAATTCATATCCATGAGAGATATGAGGACT 1050
 Db 1009 CTGATGCTGATATTTTGGGAGATTTGGAATTCATATCCATGAGAGATATGAGGACT 1068

RESULT 12
 CD831928 554 bp mRNA linear EST 10-JUL-2003
 LOCUS CD831928
 DEFINITION BM40.061P15F011227 BM40 Brassica napus cDNA clone BM40061P15, mRNA
 sequence.
 CD831928
 VERSION CD831928.1 GI:32513868
 KEYWORDS
 SOURCE EST.
 ORGANISM Brassica napus (rape)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE
1 (bases 1 to 554)
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosida II; Brassicales; Brassicaceae; Brassica.
AUTHORS
Genoplatane.
TITLE
Unpublished (2003)
JOURNAL
Contact: Genoplatane
COMMENT
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplatane' (<http://www.genoplatane.com> and <http://genoplatane-info.infobiogen.fr>).
location/Qualifiers

FEATURES
source
1..554
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jet Neuf"
/db_xref="taxon:3708"
/clone="BN40061P15"
/issue_type="seed"
/clone_id="BN40"

ORIGIN

Query Match 41.2%; Score 455.4; DB 6; Length 554;
Best Local Similarity 97.7%; Pred. No. 8,4e-129;
Matches 462; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 ATGGCGGTGAAGAAAGGCTAGCGAGATGTTTTCGAAGAGCTTGATCGAGAGCTTTCACAGA 60
DB 81 ATGGCGGTGAAGAAAGGCTAGCGAGATGTTTTCGAAGAGCTTGATCGAGAGCTTTCACAGA 140
QY 61 TGGGAGATGATGAAGACGAGCGGGGTGAGCTTCAGTACATGATGAGATTGGTTCACACT 120
DB 141 TGGGAGATGATGAAGACGAGCGGGGTGAGCTTCAGTACATGATGAGATTGGTTCACACT 200
QY 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAAGAGAGCTTCGATTCGG 180
DB 201 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAAGAGAGCTTCGATTCGG 260
QY 181 ATGCGGAGGCGTGGATCGAATCTGAGACGCTGCTTATGCGCTCTCTGAGAAACCTGCC 240
DB 261 ATGCGGAGGCGTGGATCGAATCTGAGACGCTGCTTATGCGCTCTCTGAGAAACCTGCC 320
QY 241 GTCTTGAAGGTAAAGATTGGTATGTGAGATCATTCAGGACATGAGAGCTTTCCTGAG 300
DB 321 GTCTTGAAGGTAAAGATTGGTATGTGAGATCATTCAGGACATGAGAGCTTTCCTGAG 380
QY 301 ATCAAGATATCTGCTGATGAGAAAGATTCAACAGATGATCAAGGCTGTTAAAGTAAG 360
DB 381 ATCAAGATATCTGCTGATGAGAAAGATTCAACAGATGATCAAGGCTGTTAAAGTAAG 440
QY 361 CACAACAACTGTTTCCATGATGCTCTGGGTGTGAACCACTGAAAGAAAGTAAG 420
DB 441 CACAACAACTGTTTCCATGATGCTCTGGGTGTGAACCACTGAAAGAAAGTAAG 500
QY 421 CTCTACGAAAGCTTGAATGAGATTCATCAGTTTCTTGAATGCTTCTACTTCTGTC 473
DB 501 CTCTACGAAAGCTTGAATGAGATTCATCAGTTTCTTGAATGCTTCTACTTCTGTC 553

RESULT 13

BX839466 784 bp mRNA linear EST 11-FEB-2004
LOCUS BX839466 Arabidopsis thaliana Adult vegetative tissue Col-0
DEFINITION Arabidopsis thaliana cDNA clone G517L534ZB07 5PRIM, mRNA sequence.
ACCESSION BX839466
VERSION BX839466.1 GI:42533549
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE
1 (bases 1 to 784)
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosida II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS
Castelli V., Aury J.M., Jailion O., Wincker P., Clepet C., Menard M., Cruaud C., Quetier F., Scarpelli C., Schachter V., Temple G., Caboche M., Weissenbach J. and Salanoubat M.
TITLE
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL
Unpublished (2004)
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Creteil, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction; Temple G.
Genoscope members carried out sequencing and annotation; Castelli V., Aury J.M., Jailion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences).
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.
location/Qualifiers

FEATURES

source

ORIGIN

Query Match 40.8%; Score 450.2; DB 5; Length 784;
Best Local Similarity 82.5%; Pred. No. 3,9e-127;
Matches 565; Conservative 0; Mismatches 113; Indels 7; Gaps 4;

QY 1 ATGGCGGTGAAGAAAGGCTAGCGAGATGTTTTCGAAGAGCTTGATCGAGAGCTTTCACAGA 60
DB 104 ATGGCGGTGAAGAAAGGCTAGCGAGATGTTTTCGAAGAGCTTGATCGAGAGCTTTCACAGA 163
QY 61 TGGGAGATGATGAAGACGAGCGGGGTGAGCTTCAGTACATGATGAGATTGGTTCACACT 120
DB 164 TGGGAGATGATGAAGACGAGCGGGGTGAGCTTCAGTACATGATGAGATTGGTTCACAA 223
QY 121 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAAGAGAGCTTCGATTCGG 180
DB 224 CCCACTGAGAGAAACCTTCTGATCTCGGCGAGTTTCTTCAAGAGAGCTTCGATTCGG 283
QY 181 ATGCGGAGCGTGGATCGAATCTGAGACGCTGCTTATGAGCTCTGAGAAACCTGTC 239
DB 284 ATGCGGAGCGTGGATCGAATCTGAGACGCTGCTTATGAGCTCTGAGAAACCTGTC 343
QY 240 CGTCTTGAAGGTAAAGATTGGTATGTGAGATCATTCAGGACATGAGAGCTTTCCTGA 299
DB 344 CGTCTTGAAGGTAAAGATTGGTATGTGAGATCATTCAGGACATGAGAGCTTTCCTGA 403
QY 300 GATCAAGATATCTGCTGATGAGAAAGATTCAAC-AGATGATCAAGGCTGTTAAAGTAA 358
DB 404 GATTAAGATTCGGGTGAGAGAAAGTTTCACTCAAGATGATTAAGGCTGCAAGTAA 463
QY 359 GGCACAACAACTGTTTCCATGATGCTCTGGGTGTGAACCACTGAAAGAAAGATGA 418
DB 464 GGCACAACAACTGTTTCCATGATGCTCTGGGTGTGAACCACTGAAAGATGA 523
QY 419 AACTTACGAAAGCTTGAATGAGATTCATCAGTTTCTTGAATGCTTCTACTTCTGCTGA 478
DB 524 TTCT---GGAAATCTTGAATGAGATTCATCAGTTTCTTGAATGCTTCTACTTCTGCTGA 579

QY 479 TAGGATCCGATGCTTA-TCGGGACAGATGTTGAGTTCATATCAAAACCACCACTT 537
 DB 580 TCGGATCCGGGTGCTTATTTGGGACAGCTTGATTCATATCAATCCACCGCTT 639
 QY 538 CACACAGTGGGTATACATACACCAAGATGTCCTATGAGAGGTGGCAAGATGCTAT 597
 DB 640 CATACAGTGGGTATATACACCAAGATGTCCTATGAGAGGTGGCAAGATGCTAT 699
 QY 598 GAGATGCAAGTTCGATTTGTTTCAAGATGTTGCTTCCGAGATTAACATATAT 657
 DB 700 GAGATGTCGAGCAATTTTTCGAGAGTGGGTGCTCCGGATTAACATATAT 759
 QY 658 GCGATCCAAAGTTCACTTTCCGT 682
 DB 760 GGGATTCCAATTTTACCTTCCGT 784

RESULT 14
CK208011

LOCUS 1090 bp mRNA linear EST 08-DEC-2003
 DEFINITION FGAS019688 Triticum aestivum FGAS: Library 5 GATE 7 Triticum

ACCESSION CK208011
 VERSION CK208011
 KEYWORDS GI:39570401

SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum

REFERENCE
AUTHORS

Allard, F., Crosby, W. L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,
 Genswein, B., Graf, R., Gulick, P., Hrycan, L. D., Laroche, A.,
 Links, M. G., McCarthy, E. L., Monroy, A., Muzak, I., Nilsson, D.,
 Peniket, C., Roach, J. L. and Sarhan, F.,
 Functional Genomics of Abiotic Stress in Wheat and Canola Crops
 Unpublished (2003)
 Contact: Wm L Crosby
 Bioinformatics
 University of Saskatchewan, Department of Computer Science
 1C101 Engineering Building, 57 Campus Drive, Saskatoon,
 Saskatchewan, S7N 5A9, Canada
 Tel: 306 966 1769
 Fax: 306 966 2033
 Email: fgas_estecs.usask.ca

TITLE
JOURNAL
COMMENT

This sequence is the direct result of the Base calling software
 Phred (default parameters). It is the raw base call. To aid in the
 identification of the high quality insert the software Lucy
 (default parameters) has been run on this sequence. Lucy identified
 the region [25, 811].
 Plate: LSB012 row: K column: 21.

FEATURES
SOURCE

Location/Qualifiers
 1..1090
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
 /note="Vector: pCMV.SPORT6; crown and developmental stages
 of spike formation in wheat cultivar Norstar. 4 mRNA
 populations were combined before constructing the library.
 The first mRNA population is from 1cm crown sections after
 30 days of cold acclimation. The second is from 1cm crown
 sections after 11 days of deacclimation (before
 deacclimation plants were fully vernalized for 49 days).
 The third is from different developmental stages of spike
 formation (5 to 50mm) that still have not emerged from the
 leaf (dissection required). The last is from different
 developmental stages of spike and seed formation after
 having emerged from the leaf (visible). First strand
 synthesis in this library was done in the presence of
 methylated dCTP thereby protecting from internal cleavage
 with NotI."

ORIGIN

Query Match 40.0%; Score 441.6; DB 7; Length 1090;
 Best Local Similarity 68.9%; Pred. No. 2e-124;
 Matches 639; Conservative 0; Mismatches 276; Indels 13; Gaps 2;

QY 3 GCGGTGAGAGAGGCTTACGAGATGTTTGAAGACCTTGATGAGAGCTTCAAGATG 62
 DB 100 GACGGGGAGATGGCCGCGAGCCGGTGGCGCGCGCGCGCGCGCGCGCGCGCG 159
 QY 63 GGGATGATTAAGACGAGCGGCGTGAAGCTTCAAGTATCATATGAGATTGCTCACTC 122
 DB 160 GGGCGGATGAGGCAACAGATGTCAGCTGCGCTTACATGAGAGATTGGGGCGCGCC 219
 QY 123 CACTGAGAGAACCTTCTGATCTCGGCGAGTTTCTTCAAGAGGCTTCCGATTCCGAT 182
 DB 220 CACGAGCGGCAACTCTGCTGCTCCCGGAGTTCTTCCGACAGAGAGCTCCCATCCGAT 279
 QY 183 CCGGAGCGTGGCATGAACTGAGAGCGTCCCTTATGAGCTTCTGAGAACTGCGCT 242
 DB 280 CCGCAGCGCGCGCTGACCTGACCTCCCTCCCTTCCGCTTCCATTAAGCCGCGCT 339
 QY 243 CTGAGAGTAAAGATTGATGATGAGATTCATTCAGGACATGAGAGCTTCTGAGAT 302
 DB 340 CTTCAAGTGAAGATGTTGTTTACCTGATTCGTTCCGATCCTGCTTCCGAGAT 399
 QY 303 CAGGATTAAGTCTGATGAGAGAGTTCAACAGATGATCAAGGCTTAAAGTAAAGCA 362
 DB 400 GAGCAACGAGAGATGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 459
 QY 363 CACAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 412
 DB 460 CACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 519
 QY 413 -GAATGAACCTTCAAGAAAGCTTGAATGATGATGATGATGATGATGATGATGAT 470
 DB 520 TACAAAGGATTCCTCTCGAATTAAGATGATGATGATGATGATGATGATGATGAT 579
 QY 471 GTCTGTATGAGATCCGATGCTTATGAGGAGATGATGATGATGATGATGATGATGAT 530
 DB 580 GTCAAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 639
 QY 531 ACCACTTCAACAGGAGTATACATACACCAAGATGCTCTTGAAGGAGGCAAGAA 590
 DB 640 AGAGCCTGCGATATAGGCTCATTAACAAAGATGCTCCCATGCTGAGTCCGACT 699
 QY 591 TGCTGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 650
 DB 700 TGCTGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 759
 QY 651 CATATATGCGATCAAGATTCATCTTCCGATGCTCCAGACCATTTGATCTTATGAT 710
 DB 760 CATATATGCGATCAAGATTCATCTTCCGATGCTCCAGACCATTTGATCTTATGAT 819
 QY 711 GTATGATTAAGTCAAGATTCCTCCGATGCTCCAGACCATTTGATCTTATGAT 770
 DB 820 GTTGAATGATGAAGATTCCTCCGATGCTCCAGACCATTTGATCTTATGAT 879
 QY 771 GGTGACCAACCATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 830
 DB 880 ACATGCACTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 939
 QY 831 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 890
 DB 940 TAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 998
 QY 891 CAGATGCAAGAAACCATTTGAAGA 918
 DB 999 CCGTACGAGAAACCATTTGAAGA 1026

RESULT 15
 CN892416

LOCUS CN892416 669 bp mRNA linear EST 04-JUN-2004
 DEFINITION 010603AA005454HT (AXXA) Royal Gala 126 DAFB fruit core Malus x
 domestic cDNA clone AXA005454, mRNA sequence.
 ACCESSION CN892416
 VERSION CN892416.1 GI:48278658
 KEYWORDS EST.
 SOURCE Malus x domestica (cultivated apple)
 ORGANISM Malus x domestica
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids 1; Rosales; Rosaceae; Maloideae; Malus.
 REFERENCE 1 (bases 1 to 669)
 AUTHORS Beuning, L., Bowen, J., Crowhurst, R., Gleave, A., Janssen, B.,
 McCartney, S., Newcomb, R., Ross, G., Snowden, K., Walton, E. and Yauk, Y.
 TITLE HortResearch Apple EST Project
 JOURNAL unpublished (2004)
 COMMENT Contact: Gleave, A.
 Sequencing Facility
 The Horticulture and Food Research Institute of New Zealand Ltd
 120 Mc Albert Rd, Mc Albert, Auckland, New Zealand
 Tel: 00 64 09 815 4200
 Fax: 00 64 09 815 4201
 Email: eest@hortresearch.co.nz.
 FEATURES
 source
 1..669
 Location/Qualifiers
 /organism="Malus x domestica"
 /mol_type="mRNA"
 /db_xref="taxon:3750"
 /clone="AXXA005454"
 /issue_type="fruit core"
 /dev_stage="126 days after full bloom"
 /clone_id="(AXXA) Royal Gala 126 DAFB fruit core"
 /note="Vector: pBluescript SK(-); library sequenced by
 Genesis Research & Development"

ORIGIN

Query Match 37.0%; Score 409; DB 7; Length 669;
 Best Local Similarity 78.3%; Pred. No. 1.9e-114;
 Matches 490; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 1 ATGCGCGTGAAGAGGCTTACGAGATGTTTCCAGAGCTTGATCGAGAGCTTACACA 60
 DB 44 ATGCGCGTGAAGAGGCGAGCGAGTCTTCTCGAAGAGCTTGATCGAGAGCTGACAGA 103
 QY 61 TGGGGATGATGAGAGCAGAGCGGCGTGAAGCTGATGATGATGATGATGATGATGAT 120
 DB 104 TGGGGATGATGAGAGCAGAGCGGCGTGAAGCTGATGATGATGATGATGATGATGAT 163
 QY 121 CCCACTGAGAGAACTTCTGATCTCGCGCAGTTTCTTCAAGAGCTTCCGATTCCG 180
 DB 164 CCGACGAGCGGAGATTCTATATCTCGCGCAGTTTCTTCAAGAGCTTCCGATTCCG 223
 QY 181 ATGCGAGGCGTGGATGAGCTGAGAGCGTCTTATGAGCTCTTGAAGAACTGCG 240
 DB 224 ATGCGAGGCGGCGCAATTGAGCTCGAGAGCTCCCTTATGAGCTTGTGAGAACTGCG 283
 QY 241 GTCTTGAAGGTAGAGATTGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 DB 284 GTTTTGAAGGTAGAGATTGATGATGATGATGATGATGATGATGATGATGATGATGAT 343
 QY 301 ATCAAGGATAGCTGATGAGAGAGTTCACACAGATGATCAAGGCTGTTAAAGTAAG 360
 DB 344 ATCAAGGATAGCTGATGAGAGAGTTCACACAGATGATCAAGGCTGTTAAAGTAAG 403
 QY 361 CACAAACAGCTGTTCCCATGATGAGCTTGGGTGTGAACCACTGAAGAAAGGATGAA 420
 DB 404 CACAAACAGCTGTTCCCATGATGAGCTTGGGTGTGAACCACTGAAGAAAGGATGAA 463
 QY 421 CTCTAGAGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 DB 464 CTCTAGAGAGATCTTGAATGAAATTCACCAATTTTGAATGATGATGATGATGATGAT 523
 QY 481 GGGATCGATGCTTATCGGCGAGCATGTTGATGATGATGATGATGATGATGATGATGAT 540

DB 524 GGGATCGATGCTTATGCGCAGCATGTTGATGATGATGATGATGATGATGATGATGATGAT 583
 QY 541 ACAGTGGTTACATACACACCAAGATGCTTCCATAGAGAGTGGCAAGGATGATGATGATGAT 600
 DB 584 TGTGTGGTTATATGATACAAAGATGCTTCCATAGAGAGTGGCAAGGATGATGATGATGAT 643
 QY 601 GATCAAGGTCGATTTGTTTCAAGAGA 626
 DB 644 GAGCCCGTGCATGTCCTGCGCTGA 669

Search completed: April 12, 2005, 08:04:50
 Job time : 4258 secs

This Page Blank (uspto)